NASA-CR-172725
19830019754

The Telecommunications and Data Acquisition Progress Report 42-73

January – March 1983

E. C. Posner
Editor

May 15, 1983

NASA
National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The Telecommunications and Data Acquisition Progress Report 42-73

January - March 1983

E. C. Posner
Editor

May 15, 1983

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

Preface

This quarterly publication provides archival reports on developments in programs managed by JPL's Office of Telecommunications and Data Acquisition (TDA). In space communications, radio navigation, radio science, and ground-based radio astronomy, it reports on activities of the Deep Space Network (DSN) and its associated Ground Communications Facility (GCF) in planning, in supporting research and technology, in implementation, and in operations. Also included is TDA-funded activity at JPL on data and information systems and reimbursable DSN work performed for other space agencies through NASA. The preceding work is all performed for NASA's Office of Space Tracking and Data Systems (OSTDS).

In geodynamics, the publication reports on the application of radio interferometry at microwave frequencies for geodynamic measurements. In the search for extraterrestrial intelligence (SETI), it reports on implementation and operations for searching the microwave spectrum. The latter two programs are performed for NASA's Office of Space Science and Applications (OSSA).

Finally, tasks funded under the JPL Director's Discretionary Fund and the Caltech President's Fund which involve the TDA Office are included.

This and each succeeding issue of the TDA Progress Report will present material in some, but not necessarily all, of the following categories:

OSTDS Tasks:

- DSN Advanced Systems
  - Tracking and Ground-Based Navigation
  - Communications, Spacecraft-Ground
  - Station Control and System Technology
  - Network Data Processing and Productivity
- DSN Systems Implementation
  - Capabilities for New Projects
  - Networks Consolidation Project
  - New Initiatives
  - Network Sustaining
- DSN Operations
  - Network Operations and Operations Support
  - Mission Interface and Support
  - TDA Program Management and Analysis
- GCF Operations and Implementation
- Data and Information Systems

OSSA Tasks:

- Search for Extraterrestrial Intelligence
- Geodynamics
  - Geodetic Instrument Development
  - Geodynamic Science

Discretionary Funded Tasks

This Page Intentionally Left Blank

Contents

OSTDS TASKS
DSN Advanced Systems
COMMUNICATIONS, SPACECRAFT-GROUND

DSN Systems Implementation
CAPABILITIES FOR NEW PROJECTS

NETWORKS CONSOLIDATION PROJECT

NETWORK SUSTAINING

DSN Operations

NETWORK OPERATIONS AND OPERATIONS SUPPORT

TDA PROGRAM MANAGEMENT AND ANALYSIS

GCF Operations and Implementation

CCIR SUPPORT

The JPL 1.5-Meter Clear Aperture Antenna With 84.5 Percent Efficiency

A. G. Cha
Radio Frequency and Microwave Subsystems Section

This paper details the theoretical and experimental study of a 1.5-meter offset dual-shaped reflector at 31.4 GHz. An efficiency of 84.5 percent, a likely new record for reflector antennas, was ascertained through careful measurements. For larger low noise reflector systems, a 2- to 3-dB improvement in G/T performance over the state-of-the-art ultra-low-noise ground stations and 90 percent or better aperture efficiency now appear feasible.

I. Introduction

Present ground station antennas are mostly of the symmetric dual-reflector type. In the last few years, an investigation has been made of alternative ground station antenna designs which have a clear aperture (no subreflector or strut blockage) and shaped reflector surfaces. The new configurations are shown to have a potential 2- to 3-dB performance enhancement in gain over noise temperature ratio (G/T), for low noise (20 K total) systems, compared to the present symmetric reflector stations as well as an aperture efficiency in excess of 90 percent (Ref. 1). Much progress in the synthesis (Refs. 2–8), analysis, and design approaches (Refs. 9, 10) of this new type of antenna has been accomplished in recent years. A small demonstration model with a main reflector diameter of 1.5 m was designed, fabricated, and tested at 31.4 GHz to confirm the theoretical synthesis and analysis procedures.

The antenna consists of a 157.1λ main reflector, 47.1λ subreflector and a 5.1λ aperture feedhorn. Both reflectors are shaped to provide quasi-uniform aperture illumination with a degree of spillover control. The configuration provides an unblocked main reflector by use of an offset feed geometry. An analysis of expected performance yields a theoretical 86.5 percent aperture efficiency while our measurement technique proves an actual 84.5 percent. The estimated measurement tolerance is presently ±3 percent (about ±0.15 dB) on a high confidence (approximately 2σ) basis. To the best of our knowledge, there is no other reflector antenna having such high aperture efficiency (Refs. 11, 12). Figure 1 shows the assembly which was constructed as a proof-of-design model. Economics dictated a somewhat smaller assembly than desired; we project on the same theoretical basis that a larger D/λ subreflector, to minimize diffraction, will yield aperture efficiency beyond 90 percent.

We define aperture efficiency in the usual way to include spillovers, illumination, phase and cross polarization efficiencies of the projected circular aperture, although we do not include feedhorn dissipation since in principle that is subject to good control.

In this paper, we present an overview of the synthesis, analysis, and design techniques used in realizing the 1.5-m model. While the synthesis defining the reflector surface geometry is based on geometric optics, the analysis procedure utilized is based on physical optics, which takes into account the (diffraction) rear spillover loss beyond the main reflector edge and the vector nature of the field; i.e., both principal and cross polarized field components are considered. An extensive experimental program designed to verify the theoretical predictions is also described. Excellent correlation between theory and experiment has been achieved in both secondary gain and pattern characteristics and subreflector diffraction characteristics. The development herein is applicable to both maximum gain and maximum G/T designs. The 1.5-m model is based on a maximum gain design.

II. Synthesis

The basic synthesis used in this paper was discussed in great length in Refs. 2–4. The approach yields an approximate synthesis solution based on ray optics principles, i.e., constant pathlength for each ray, Snell's law on each reflector surface, and conservation of energy. In Fig. 2, (r, θ, ϕ) defines a spherical coordinate system with the feed phase center as the origin and (ρ, ψ, z) defines a cylindrical coordinate system with the main reflector center as the origin. The path length along the trajectory to a plane located at $z = D$ is

$$r + S + D - z = \text{constant} = K' \tag{1}$$

where S is the path length between the reflectors.

There are four Snell's law equations—two for each reflector—because of the double curvature of each reflector. We determine these equations by applying Fermat's principle of minimum length for the optical path. This leads to the following equations on the main and subreflectors respectively.

$$z_\rho - S_\rho = 0 \text{ and } z_\psi - S_\psi = 0 \tag{2}$$

$$r_\theta + S_\theta = 0 \text{ and } r_\phi + S_\phi = 0 \tag{3}$$

Let us examine the energy relationship. Differentially, we require conservation of energy, as expressed by

$$V_c V(\rho, \psi)\, \rho d\rho d\psi \equiv I(\theta, \phi) \sin\theta\, d\theta d\phi \tag{4}$$

where V_c is a proportionality constant to be determined. While the exact solution to the simultaneous algebraic and partial differential equations in (1) to (4) is unknown, a two-step approach to obtain an approximate solution was developed in Ref. 2. In the first step, a mapping equation

$$\psi = \phi \tag{5}$$

is introduced. This significantly simplifies the mathematics by allowing Eqs. (1) to (4) to be reduced to two simultaneous ordinary differential equations, which are then solved numerically. Since Eq. (5) is only approximately correct, the solution so obtained is only approximate. In the second step, only the subreflector obtained in the first step is retained and a new matching main reflector is designed by imposing the constant path length requirement. This results in an approximate synthesis solution where the output reflector aperture amplitude distribution has a small deviation from the prescribed distribution although the aperture phase distribution always follows the prescribed one exactly (generally a uniform phase distribution is desired). Another approximation accepted in the synthesis is that the output main reflector periphery cannot be prescribed. In spite of these limitations, the extensive cases obtained in Refs. 2 and 4 illustrate the effectiveness and flexibility of the technique in obtaining many practical solutions. The flexibility demonstrated has included specifying the main reflector aperture distribution, the main to subreflector diameter ratio, and the quasi f/D ratio. In all cases, the resultant main reflector aperture is within 1% of being circular. For practical considerations, the basic synthesis has seen two extensions to improve its numerical efficiency (Ref. 3) and to ensure desirable RF characteristics (Ref 4).

As compared to circularly symmetric designs, a large quantity of data for the reflector surfaces is required in an offset design. An extremely efficient synthesis algorithm has been developed for this purpose. It is based on the requirement that the reflector surfaces must be smooth. Thus the bulk of the data defining the surfaces can be obtained through an interpolation process making use of one-dimensional cubic spline functions. This "efficient" version is discussed in Ref. 3. Due to the approximate nature of the synthesis, the solutions obtained sometimes show a brightly lit boundary in the main reflector aperture based on geometric optics and physical optics analysis. See, for example, Fig. 1 of Ref. 1. There was some concern that such designs could have correspondingly high rear (diffraction) spillover loss passing the main reflector, a very detrimental effect causing excessive antenna noise temperature in low noise systems. In Ref. 4, a strategy was adopted for controlling the diffraction spillover by synthesizing an illumination which is nearly uniform in the aperture except near the edge where the illumination falls off sharply. The aperture power distribution specified is

$$V(\rho, \psi) = (1 - \rho^2)^p \tag{6}$$

where $0 \leqslant p \leqslant 1$. Note for $p = 0$ we have the conventional optically uniform illumination case. For small positive p values, one has a slightly tapered illumination design with reduced rear spillover loss but at the expense of a small reduction in illumination efficiency. This was verified by diffraction analysis (Ref. 4).

III. Performance Analysis and Optimization

The analysis serves many important purposes. It enables one to (1) determine the performance of the reflector, (2) select the best design among a number of cases, and (3) better understand the physics and seek design improvements. Common performance parameters include the gain (efficiency), G/T and far field sidelobe and polarization characteristics. A block diagram of the reflector performance analysis procedure is shown in Fig. 3. Each efficiency term represents a loss mechanism within the antenna system. The overall antenna efficiency η_T is

$$\eta_T = \eta_{FS}\,\eta_{RS}\,\eta_p\,\eta_X\,\eta_I\,\eta_{RMS}\,\eta_D \tag{7}$$

The terms η_{RMS} and η_D are losses due to surface RMS error and dissipation in the antenna and feed. These are not included in the diffraction and efficiency analysis. Notably missing from Eq. (7) is the blockage efficiency η_B as a clear aperture design has been assumed. Note, in Fig. 3, the efficiency η_T can only be evaluated after all such intermediate computations as subreflector diffraction patterns and complex aperture excitations have been calculated. This observation applies to all other performance parameters, e.g., antenna noise temperature and far field radiation characteristics.

The procedure in Fig. 3 may be performed repeatedly to analyze different designs and arrive at an optimal selection. As noted, the principal concern is the potential degrading effects of rear spillover loss. The design tactic used in this paper consists of determining the aperture illumination taper, or the value of p in Eq. (6), based on either the optimal gain or the optimal G/T criterion. An alternative approach is adding a noise shielding flange to the main reflector of the $p = 0$ design to redirect most of the rear spillover energy into the sky (Refs. 13, 14). Figure 4 shows the physical optics computed diffraction patterns of two 20λ shaped subreflectors, taken from Ref. 4. The two designs have $p = 0$ and 0.25 respectively but otherwise identical geometric parameters and feed. Although one expects a nominal loss in illumination efficiency for the $p = 0.25$ case, we observe that both the diffraction spillover and the depolarization losses have been reduced. There is thus a tradeoff between the illumination and rear spillover efficiencies (and antenna noise temperature).

From the subreflector far field diffraction pattern, one can proceed to compute the complex, vector incident field at the main reflector and the rear spillover efficiency. The rear spillover efficiency is computed from standard power integration routines that determine the percentage of subreflector diffracted power intercepted by the main reflector. The complex vector incident field at the main reflector is found using a "parallax correction" adapted from that developed for the symmetric dual-shaped reflector system (Refs. 13, 15). This correction is shown in Fig. 5 for the asymmetric dual-reflector system. In computing the incident field at a point P on the main reflector (Fig. 5) using far field physical optics (FFPO), a phase reference point O is defined. The actual ray path based on geometric optics is from the feed phase center F to the point Q on the subreflector and then the point P on the main reflector. It is easy to see that the derivation in Ref. 15 for the symmetric dual-reflector case is valid for the present asymmetric reflector case. The incident field at point P is given by the FFPO computed field, with O as the origin, in the $\hat{R}'$ direction with an additional phase correction $e^{j\Phi}$ where

$$\begin{aligned} \Phi &= k(R - R' - \overline{00'} \cdot \hat{R}') \\ &= k(R - R\cos\angle 0P0') \\ R &= |\overline{0P}| \\ R' &= |\overline{0'P}| \end{aligned} \tag{8}$$

The parallax correction is applied to both the principal and cross polarized field component. The reflected field and aperture field are readily determined from the incident field from Snell's law. It is interesting that in the asymmetric dual-reflector case the above parallax correction corrects not only the large phase error of the incident field to the main reflector but also a significant polarization error. Without this correction, it is found that the derived aperture field contains a significant axial component normal to the aperture.

The computation of the secondary field $\bar{\mathscr{E}}$ from the aperture $\bar{E}$ and $\bar{H}$ field is a well-known procedure given in Ref. 16 by

$$\bar{\mathscr{E}}(\theta, \phi) = \frac{-jk\,e^{-jkR}}{4\pi R}\ \hat{R} \times \iint dS' \{\hat{n} \times \bar{E} - Z_0\,\hat{R} \times (\hat{n} \times \bar{H})\}\, e^{jk\bar{R}' \cdot \hat{R}} \tag{9}$$

where R, θ, ϕ are the spherical coordinate variables, $\hat{n}$ is the normal vector to the aperture and Z_0 is the free space impedance. The antenna gain and aperture efficiency are found from

the boresight field ($\theta = 0$, $\phi = 0$). It is often of interest to factor the aperture efficiency in the form given in Eq. (7). This allows a designer to see clearly areas that can or need to be improved. As mentioned earlier, the forward and rear spillover efficiencies are found by numerical power integration of the feed and subreflector radiation patterns. We now derive the polarization, illumination, and phase efficiencies in terms of the aperture field components E_x, E_y as follows. It is taken that E_y is the principal polarized component and E_x is the cross polarized component. The polarization efficiency is defined as

$$\eta_X = \frac{P_y}{P_x + P_y} \tag{10}$$

$$P_{\substack{x \\ y}} = \frac{1}{Z_0} \iint E_{\substack{x \\ y}} E^*_{\substack{x \\ y}} \, dS' \tag{11}$$

We now find the boresight field from Eq. (9) corresponding to the E_y component, noting that E_x does not contribute to the principal polarized field on the axis, by letting

$$\begin{cases} \hat{R} = \hat{n} = \hat{z} \\ \overline{E} = E_y \hat{y} \\ \overline{H} = \dfrac{E_y}{Z_0} \hat{x} \end{cases} \tag{12}$$

We obtain the boresight field as

$$\mathscr{E}_\theta = \frac{jk\, e^{-jkR}}{2\pi R} \iint E_y \, dS' \tag{13}$$

Note E_y is generally a complex function over the aperture S'. We define a "reference" axial field $E_{\theta IP}$ corresponding to an aperture distribution $|E_y|$ with uniform phase

$$\mathscr{E}_{\theta IP} = \frac{jk\, e^{-jkR}}{2\pi R} \iint |E_y| \, dS' \tag{14}$$

We now define the phase efficiency as

$$\eta_P = \frac{\mathscr{E}_\theta \mathscr{E}^*_\theta / Z_0}{\mathscr{E}_{\theta IP} \mathscr{E}^*_{\theta IP} / Z_0} = \frac{\left| \iint E_y \, dS' \right|^2}{\left[\iint |E_y| \, dS' \right]^2} \tag{15}$$

The gain G of the reflector is

$$G = \frac{4\pi R^2 \, (\mathscr{E}_\theta \mathscr{E}^*_\theta / Z_0)}{P_T} \tag{16}$$

where the quantity inside the parenthesis is the far field on-axis power to be evaluated at a distance R by Eq. (13) and P_T is the input power to the feedhorn. By the definition of Eqs. (10) and (11), we have

$$P_y = \eta_{FS} \, \eta_{RS} \, \eta_X \, P_T \tag{17}$$

Substituting (17) in (16) and using (15),

$$G = \eta_{FS} \, \eta_{RS} \, \eta_X \, \eta_P \, G_{IP} \tag{18}$$

$$G_{IP} = \frac{4\pi R^2 \, (\mathscr{E}_{\theta IP} \mathscr{E}^*_{\theta IP} / Z_0)}{P_y} \tag{19}$$

where G_{IP} can be thought of as the on-axis gain of an aperture energized by an uniform phase but non-uniform amplitude distribution $|E_y|$. The gain of a uniform phase and uniform amplitude aperture distribution is

$$G_0 = 4\pi S'/(\lambda^2) = \frac{1}{\pi} k^2 S' \tag{20}$$

The aperture efficiency is commonly defined as

$$\eta = G/G_0 \tag{21}$$

From (18) to (21), and the definition in Eq. (7) exclusive of η_{RMS} and η_D,

$$\eta = \eta_{FS} \, \eta_{RS} \, \eta_X \, \eta_P \, \eta_I \tag{7A}$$

we derive

$$\eta_I = \frac{G_{IP}}{G_0} = \frac{1}{G_0} \frac{k^2 \iint |E_y| \, dS'^2}{\pi \iint |E_y|^2 \, dS'} \tag{22}$$

It is seen from (22) that η_I can be thought of as the ratio of on-axis gains of two aperture distributions, that of $|E_y|$ over uniform amplitude and uniform phase.

Note that our analysis in this paper is based on far-field physical optics (FFPO), supplemented by the parallax path

length correction. Potter first pointed out that a rigorous and more accurate shaped reflector analysis should be based on near-field physical optics (Ref. 14). However, the difference between the two is most pronounced in the uniform aperture illumination case with $p = 0$ in Eq. (6). For the present slightly tapered illumination case, both the selected final design with $p = 0.5$ and its estimated performance are very close to being optimal (Ref. 10).

IV. Design of the 1.5-Meter Model

Using the synthesis and analysis techniques described in Sections II and III, a 1.5-m 31.4-GHz model was designed. The microwave design considerations and procedures were discussed in Refs. 9 and 10. The mechanical design, fabrication, alignment, and surface tolerance determination were presented in Ref. 17. The major dimensions of the model are shown in Fig. 6. The theoretical microwave performance is shown in Table 1.

A corrugated horn with 22.40-dB gain was selected as the feed because of its near perfect circularly symmetric pattern, a requirement in the present synthesis. A second significant advantage of this horn is its highly predictable pattern (Refs. 18, 19) and gain characteristics, which additionally makes this horn an excellent choice as a gain standard. This second advantage, together with the selected geometry, allows the horn to serve double functions as the feed and the gain standard in our reflector gain measurement.

V. Theory and Experiment Correlation

Three series of very carefully planned and highly accurate experiments were performed to verify the theoretically predicted characteristics. These include the feed radiation pattern measurement and gain determination, the subreflector diffraction pattern measurement and the reflector secondary pattern and gain measurements. Excellent correlations between the theory and experiments have been observed in all three series. The primary experimental goal has been the verification of the predicted 86.5 percent aperture efficiency for the reflector system. Repeated careful measurements have yielded an actual 84.5 percent.

A. Reflector Feed Pattern and Gain

The horn has a 48.36 mm aperture (1.904″, 5.07λ at 31.4 GHz) and 6.25° half flare angle. Only the fundamental (HE_{11}) mode is employed. The power pattern approximates $\cos^{84}\theta$ dependence to $\theta = 16°$, which was used in the synthesis. Figure 7 shows the theoretical and experimental horn patterns at 31.40 GHz.

For a radiator with physical circular symmetry, it has been shown that the E- and H-plane principal polarization patterns shown in Fig. 7 are sufficient to characterize its radiation pattern in the whole space and hence determine its directivity (Ref. 20). The extremely low cross polarization lobes (–48 dB) in the principal planes (Fig. 7) are not modelled by theory, giving rise to a negligible error in determining the horn directivity. The high cross polarization lobes (–36 dB) are in the diagonal planes. These are predictable from the principal polarization patterns in the two principal planes and accounted for in the directivity determination.

The technique for determining the horn directivity is by pattern integration as discussed in Refs. 21 and 22. To minimize errors in measurement and data reduction, we used three sets of experimental horn patterns about three different centers of rotation to determine the directivity at the test frequency of 31.4 GHz as suggested in Refs. 21 and 22. As a further check on accuracy, we performed this procedure at two nearby frequencies. If the directivity is G_0 at a frequency f_0, one should expect the directivity G at a slightly different frequency f to be related as

$$G_0(\text{dB}) = G(\text{dB}) - 20\log\left(\frac{f}{f_0}\right) \qquad (23)$$

given that the horn is not "gain limited", i.e., the aperture phase is sensibly uniform.

A summary of the feedhorn directivity as determined from the pattern integration technique at three frequencies is given in Tables 2 and 3. In Table 2, R is the distance between the transmitting horn and the horn under test and r is the distance between the phase center and the center of rotation of the horn under test. Note in Table 2, the directivity values at 31.4 GHz based on experimental and theoretical horn patterns are 22.403 and 22.370 dB respectively, attesting to the accuracy of the theoretical horn pattern and gain computation. Note also the very insignificant rms deviation of gain at each frequency. The three-frequency check of the gain in Table 3 is within 0.02 dB. In Ref. 21, an error analysis of the pattern integration technique is presented, concluding that the 3σ (99.7 percent confidence) accuracy of the horn gain determined is better than 0.1 dB. We note in passing that the pattern integration technique has previously been checked against other techniques at National Bureau of Standard (NBS) using an identical corrugated horn (Refs. 21, 23). The agreement between the pattern integration result and the NBS determined gain was within 0.056 dB. The estimated accuracy tolerance for the pattern integration technique is in the same ballpark as the three horn technique described by Chu and Legg (Ref. 24).

B. Subreflector Diffraction Measurements

The core of the diffraction analysis of the 1.5-m model is the physical optics (PO) computed subreflector diffraction pattern. An anechoic chamber measurement of the feed and subreflector assembly was made to ascertain the validity of this theoretical model. In Figs. 8 to 10, the reflector (or subreflector) is vertically polarized; i.e., the elevation plane is the E-plane and the azimuth plane is the H-plane. The patterns for horizontally polarized reflector (or subreflector) are nearly identical to Figs. 8 to 10. Figure 8 shows the E- and H-plane experimental and theoretical subreflector diffraction patterns. Excellent correlation between theory and experiment, down to a relative power level of between –30 to –40 dB against a relative anechoic chamber noise level of –60 dB, is observed. It is noted that the angular power distribution fully reflects what one expects to see in order to realize nearly uniform aperture illumination for the reflector system. The more constant distribution for the azimuth cut and the uneven distribution for the elevation cut are needed to compensate for the varying "space loss" of each ray in the respective planes. The subreflector diffraction measurement provides convincing support to the physical optics analysis, the success of our synthesis approach, and the final overall system high efficiency expectation.

C. Reflector Secondary Gain and Pattern Measurements

The reflector gain measurement to verify the theoretical 86.5 percent efficiency is the primary goal and also the most challenging experimental task of the 1.5-m offset dual-shaped reflector model project. With much improvement in experimental techniques over time and numerous series of measurements, we established an actual efficiency of 84.5 ±3 percent (about ±0.15 dB) on a high confidence (approximately 2σ or 95.5% probability) basis.

In Fig. 1, we show the 1.5-m-diameter main reflector (projected circular aperture normal to the wavefront), the subreflector with its focus mechanism, and the feedhorn together with a calibrated pair of precision rotary vane attenuators mounted for range tests. The feedhorn in our arrangement performs two functions: first as the system feed and secondly as the gain standard. In the gain standard mode, we physically remove the subreflector and expose the horn to a clear field of view.

This measurement technique avoids use of two horns, transmission line switching and differential insertion loss calibrations, and the accuracy depends primarily upon knowledge of the feedhorn directivity, calibration of the dual rotary vane attenuator assembly, and repeatability of the overall assembly under subreflector on/off conditions. One possible disadvantage of this technique is that the feedhorn pattern and directivity in the gain standard mode may be susceptible to interactions from the reflector surface and the baseplate, changing its value from 22.40 dBi determined in the anechoic chamber. Several experimental variations of horn placement were investigated (we call this "tire-kicking") to be certain of the validity of the gain test. In one variation, the horn is tilted up from the baseplate, which is covered by RF absorbers, in the gain standard mode. This eliminates all interactions from the metal baseplate. Mechanical flexibility is later added to the setup, allowing the feed to be moved to the front edge of the baseplate in the gain reference mode. A large sheet of RF absorber is placed behind the horn to shield it from the main reflector and the baseplate. No noticeable difference in the resulting reflector gain was observed with either arrangement, within the accuracy tolerance of the measurement. Repeatability of the measurement is excellent; the worst deviation from the average 53.12-dBi reflector gain is ±0.15 dB with the vast majority of the data points lying within ±0.1 dB of the mean value of over 100 samples taken in June and November 1982. This gain translates into an aperture efficiency of 84.5 percent.

Besides the feedhorn directivity, the other critical elements in the reflector gain measurement are two rotary vane attenuators used to compensate the nearly 30.72-dB gain difference between the reflector and the gain reference horn. These are calibrated at the test frequency of 31.40 GHz and found to have an average bias of -0.03 dB from the dial reading. This bias error has been reflected in the reported gain. Each attenuator has a calibrated range of 40 dB. Different combinations of the two are used in each measurement to provide the approximate 30.72-dB attenuation needed in bringing the reflector signal down to the level of the gain reference horn. This allows deviations of the attenuator dial setting from true values to be treated on a statistical basis. The dial is readable to a resolution of 0.025 dB. Finally, one may note the detector is operated at a constant level in this test; nonlinearities are eliminated. The antenna range used is of superlative quality with a very deep valley at the midpoint. The dynamic range available for pattern measurements is approximately 75 dB below the system peak gain of +53.12 dBi.

Figure 9 shows the far field patterns of the reflector. The main beam shapes in the two orthogonal planes are nearly identical down to -25 dB. The beamwidth is approximately 0.40 deg and the first sidelobe level is about -19 dB, all in excellent agreement with theoretical expectations. The cross polarization lobes are in the azimuth plane and are about -28 dB down from the peak of the principal polarization main lobe. The antenna also has a very clean wide angle sidelobe envelope except in the back of the main reflector (140 to 180 deg, Fig. 10), where the main reflector support structure is

present. The lobes at approximately ±18 deg are due to combined feedhorn spillover and subreflector edge diffraction effects.

VI. Conclusions and Discussions

The synthesis, analysis, and design procedures for a 1.5-m offset dual-shaped reflector antenna have been presented. Extensive experiments were performed which verified an aperture efficiency of 84.5 percent at 31.40 GHz. This is probably a new record in reflector antenna efficiency. Given the excellent correlation between theory and experiment as evidenced herein, we project that for large systems (500λ and up) 90 percent efficiency and 2 to 3 dB improvement in very low noise symmetric ground station antenna G/T performance are now possible, given high precision reflecting surfaces.

Acknowledgment

Many individuals contributed to this work and deserve full acknowledgment. For many years Hugh Fosque, Chief, Advanced Systems Office, Office of Space Tracking Data Systems at NASA Headquarters, has encouraged both development and reliable proof of high aperture efficiency reflector antennas as well as considerable and diverse other work. At JPL, the Office of Telecommunications and Data Acquisition, Advanced Systems Program, has skillfully managed and balanced systems needs and supported our work, namely Bob Clauss, Joel Smith, Robertson Stevens, and Catherine Thornton. At our technology development level, D. A. Bathker, J. A. Carpenter, W. Imbriale, V. Galindo-Israel, S. A. Rocci and W. F. Williams, as well as Raj Mittra of the University of Illinois, consultant to JPL, were responsible for important contributions. Special thanks go to H. F. Reilly who should be largely credited with obtaining the rich and highly valuable experimental information in this work.

References

1. Cha, A. G., Bathker, D., and Williams, J., "Advanced Design Concepts of Ground Station Antennas," Proceedings of 9th European Microwave Conference, Brighton, England, Sept. 1979. Also, found in *Advanced Antenna Technology*, edited by P.M.B. Clarricoats, Microwave Exhibitions and Publishers Limited, 1981.
2. Galindo-Israel, V., Mittra, R., and Cha, A., "Aperture Amplitude and Phase Control of Offset Dual Reflectors," *IEEE Trans. On Ant. Prop.*, Mar. 1979.
3. Cha, A., Galindo-Israel, V., and Mittra, R., "Efficient Design of Offset Dual Shaped Reflectors for Antenna and Beam Waveguide Applications," AP-S International Symposium, Seattle, June 1981.
4. Cha, A. G., Galindo-Israel, V., and Bathker, D. A., "Low Noise Dual Shaped Reflector Synthesis," AP-S International Symposium, Quebec, Canada, June 1980.
5. von Hoerner, S., "Minimum-Noise Maximum-Gain Telescopes and Relaxation Method for Shaped Asymmetric Surfaces," *IEEE Trans. Ant. Prop.*, May 1978.
6. Lee, J. J., Parad, L. I., and Chu, R. S., "A Shaped Offset-Fed Dual-Reflector Antenna," *IEEE Trans. Ant. Prop.*, Mar. 1979.

7. Basilaya, I. S., and Porkas, A. M., Fourth Scientific-Technical Conference of Antennas and Feeder Sections for Radio Communications, Broadcasting and Television, Moscow, Jan. 1977.

8. Westcott, B. J., Stevens, F. A., and Brickell, F., "G. O. Synthesis of Offset Dual Reflectors," *IEEE Proceedings, Part H-Microwaves, Optics and Antennas*, Feb. 1981.

9. Cha, A. G., "Design of a 1.5-m 32 GHz Clear Aperture Antenna," *TDA Progress Report 42-66*, Jet Propulsion Laboratory, Pasadena, Calif., Nov. 1981.

10. Cha, A. G., "The Synthesis, Analysis and Design of Offset Dual Shaped Reflectors," AP-S International Symposium, Albuquerque, New Mexico, May 1982.

11. Cha, A. G., and Bathker, D. A., "Preliminary Announcement of an 85 Percent Efficient Reflector Antenna," *IEEE Trans. Ant. Prop.*, Mar. 1983.

12. Cha, A. G., and Bathker, D. A., "Experimental Gain and Far Field Pattern Characteristics of a 1.5-m 85 Percent Efficient Reflector Antenna," IEEE International AP-S Symposium, Houston, Texas, May 1983.

13. Williams, W. F., "DSN 100-Meter X- and S-Band Microwave Antenna Design and Performance," Technical Report 78-65, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 1, 1978.

14. Potter, P. D., "Shaped Antenna Designs and Performance for 64-m Class DSN Antennas," *DSN Progress Report 42-20*, Jet Propulsion Laboratory, Pasadena, Calif., Apr. 15, 1974.

15. Caufield, M. F., Rusch, W. V. T., and Williams, W. F ., "Physical-Optics Analysis of a Symmetric Shaped Dual Reflector System with High Feed Tapering," IEEE International Symposium on Antennas and Propagation, Quebec, Canada, June 1980.

16. Silver, S., *Microwave Antenna Theory and Design*, Radiation Laboratory Series, No. 12, p. 161, McGraw-Hill, N. Y., 1951.

17. Carpenter, J., Rocci, S., and Chun, C. T., "The Fabrication and Surface Tolerance Measurements of the JPL Clear Aperture Microwave Antenna," *TDA Progress Report 42-71*, Jet Propulsion Laboratory, Pasadena, Calif., Nov. 15, 1982.

18. Clarricoats, P. J. B., and Saha, P. K., "Propagation and Radiation Behavior of Corrugated Feeds – Part 1. Corrugated Waveguide Feed," *Proc. IEE* (British), Vol. 118, No. 9, pp. 1167-1176, Sept. 1971.

19. Clarricoats, P. J. B., and Saha, P. K., "Propagation and Radiation Behavior of Corrugated Feeds – Part 2. Corrugated-Conical-Horn Feed," *Proc. IEE* (British), Vol. 118, No. 9, pp. 117-1186, Sept. 1971.

20. Ludwig, A., "Radiation Patterns Synthesis for Circular Horn Antennas," *IEEE Trans.*, Ap-14, No. 4, pp. 434-440, July 1966.

21. Ludwig, A., "Gain Calibration of a Horn Antenna Using Pattern Integration," Technical Report 32-1572, Jet Propulsion Laboratory, Pasadena, Calif.

22. Ludwig, A., "Gain Computations from Pattern Integration," *IEEE Trans. Ant. Prop.*, pp. 309-311, Mar. 1967.

23. Bathker, D. A., private communication.

24. Chu, T. S., and Legg, W. E., "Gain of Corrugated Conical Horns," *IEEE Trans. Ant. Prop.*, July 1982.

Table 1. 1.5-m antenna theoretical performance at 31.4 GHz, $D = 157.1\lambda$

Efficiencies	
Forward spillover	0.960
Rear spillover	0.966
Cross polarization	0.997
Phase	0.979
Illumination	0.961
Surface tolerance	0.995
Overall	0.865

Table 2. Feedhorn directivity from pattern integration (All patterns are measured except where noted otherwise)

Frequency, GHz	R, m	r, mm	Correction dB	Integrated gain, dB	Corrected gain, dB	Mean, dB	RMS deviation, dB
		101.6	−0.0727	22.272	22.199		
30.765	12.19	0	0	22.217	22.217	22.212	0.009
		−101.6	0.0721	22.147	22.219		
		141.3	−0.101	22.504	22.403		
31.40	12.19	0	0	22.402	22.402	22.403	0.001
		−115.9	0.082	22.323	22.405		
31.40 theoretical pattern	∞					22.370	
		141.3	−0.101	22.669	22.568		
32.025	12.19	0	0	22.566	22.566	22.558	0.012
		−115.9	0.082	22.459	22.541		

Table 3. Horn gain dependence on frequency

Frequency, GHz	30.765	f_0 = 31.4	32.025
G, dB From Table 2	22.212	G_0 = 22.403	22.558
$20 \log (f/f_0)$	−0.177	0	0.171
$G - 20 \log (f/f_0)$	22.389	22.403	22.387

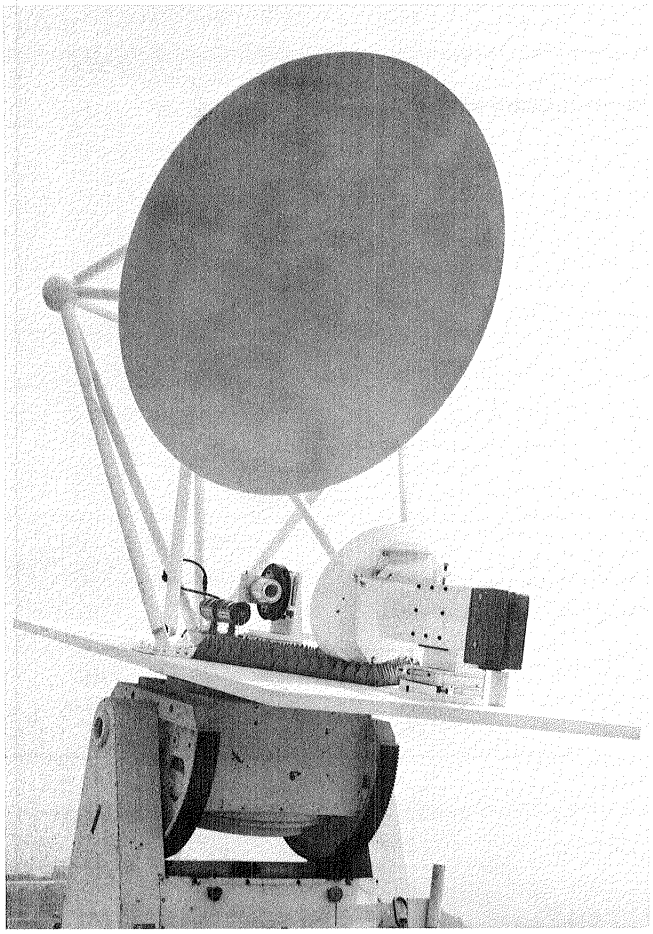


Fig. 1. 1.5-m clear aperture antenna

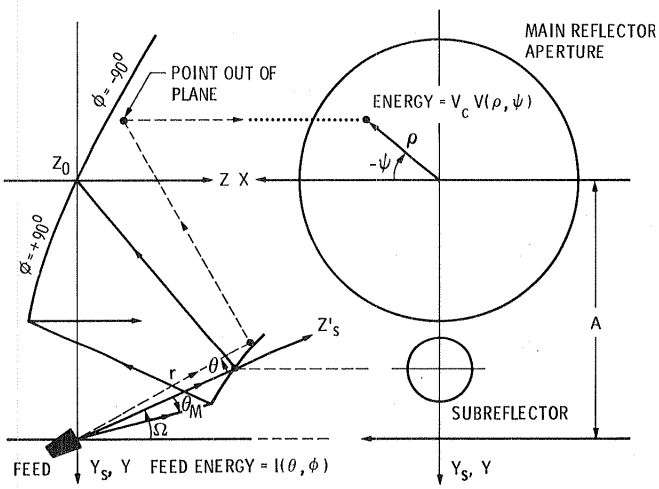


Fig. 2. Geometry of offset dual reflectors

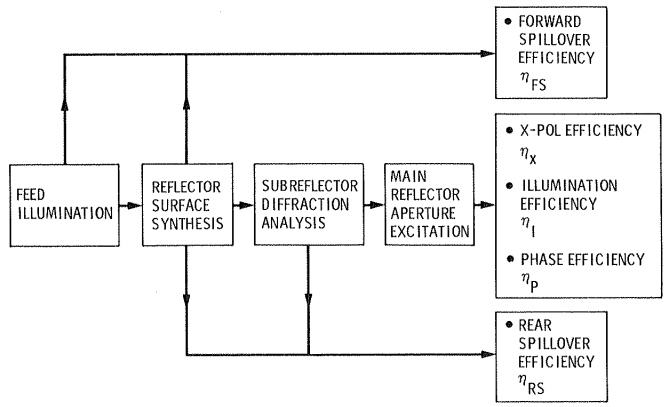


Fig. 3. Overview of reflector RF performance analysis

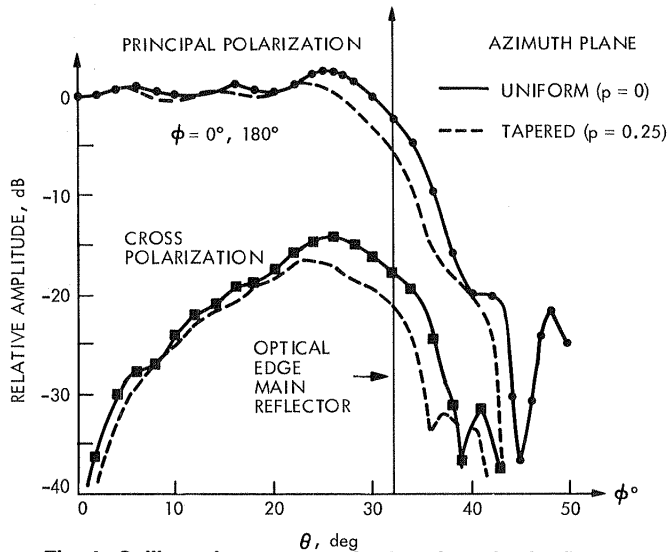


Fig. 4. Spillover improvement for 20λ shaped subreflectors

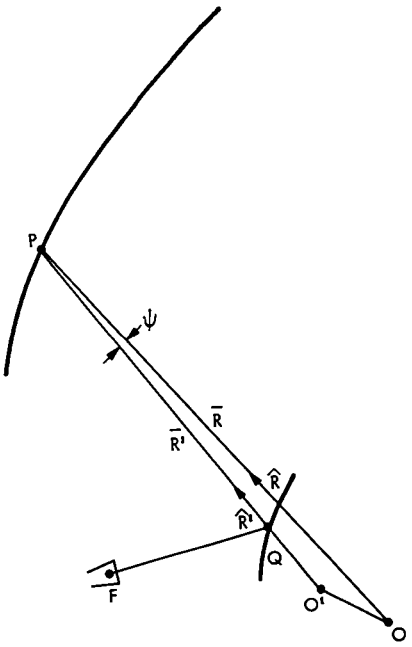


Fig. 5. Parallax correction

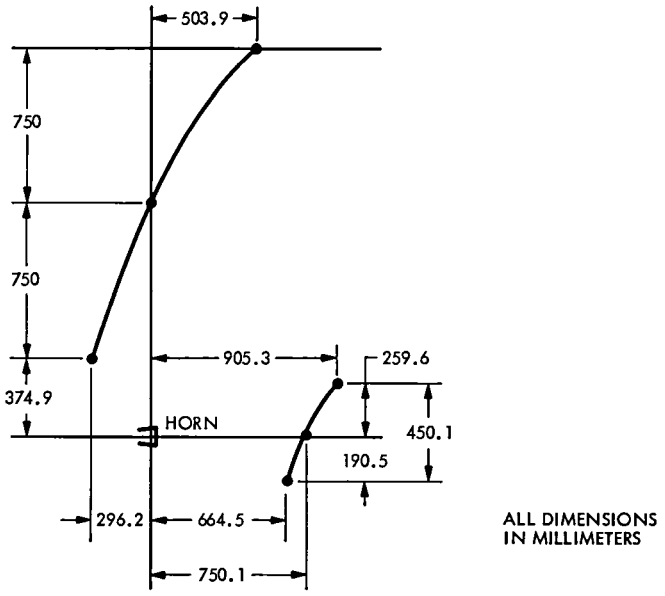


Fig. 6. Major dimensions of 1.5-m model (in mm)

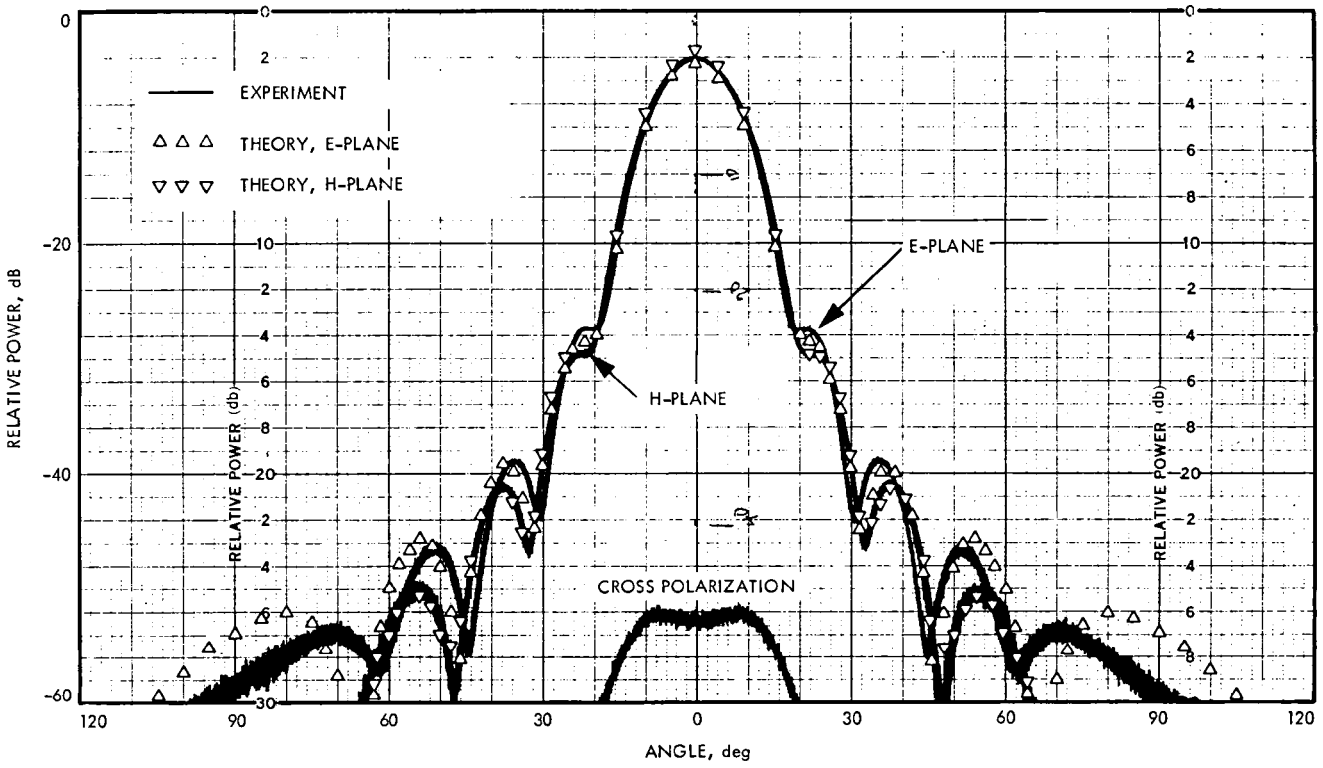


Fig. 7. Theoretical and experimental horn patterns

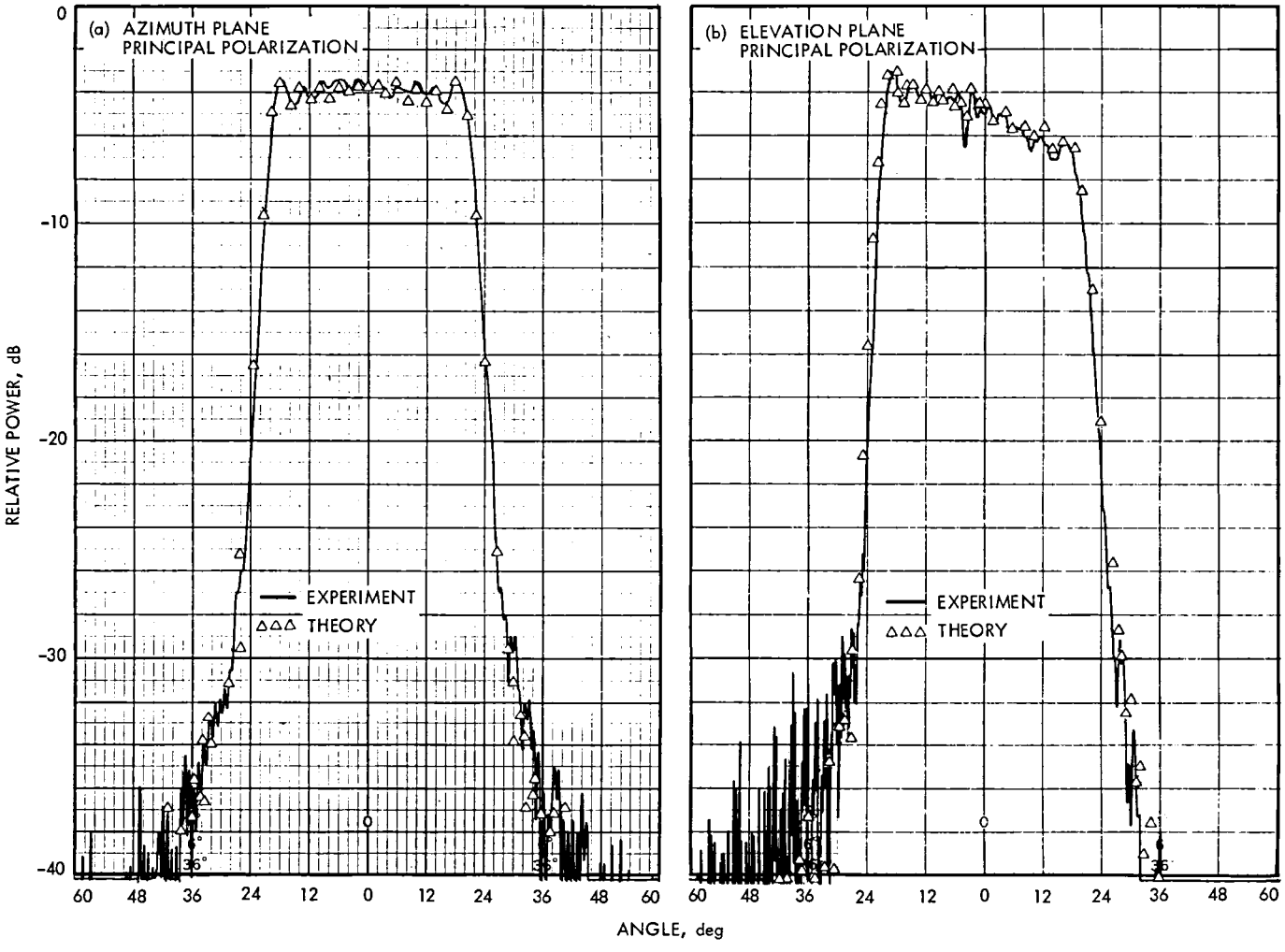


Fig. 8. Subreflector diffraction patterns

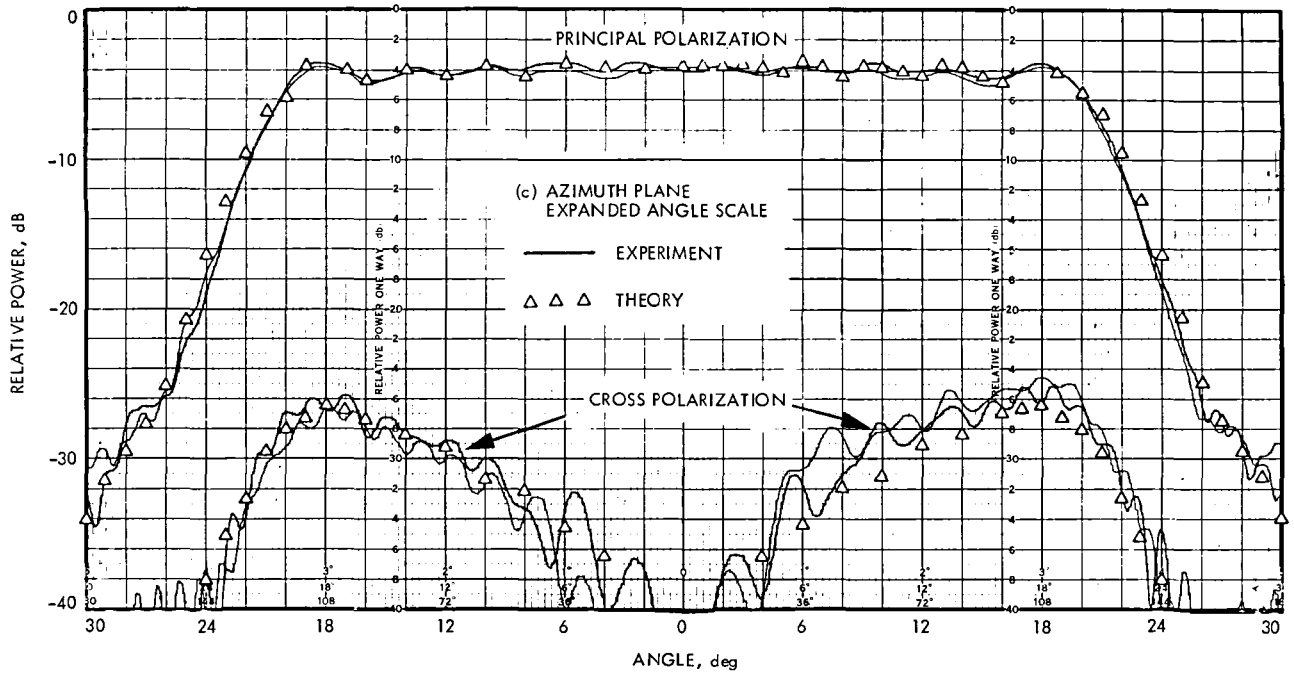


Fig. 8 (contd)

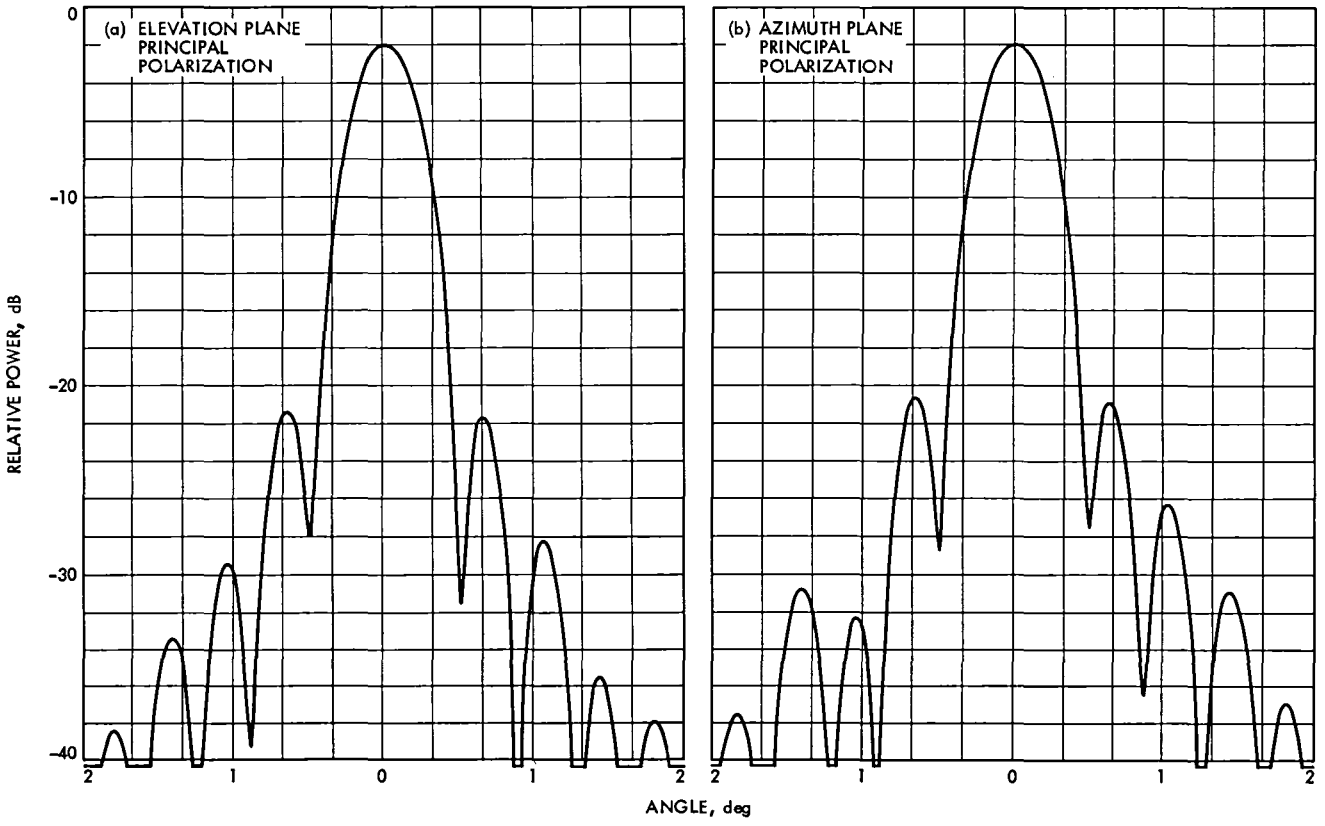


Fig. 9. Radiation patterns of 1.5-m model

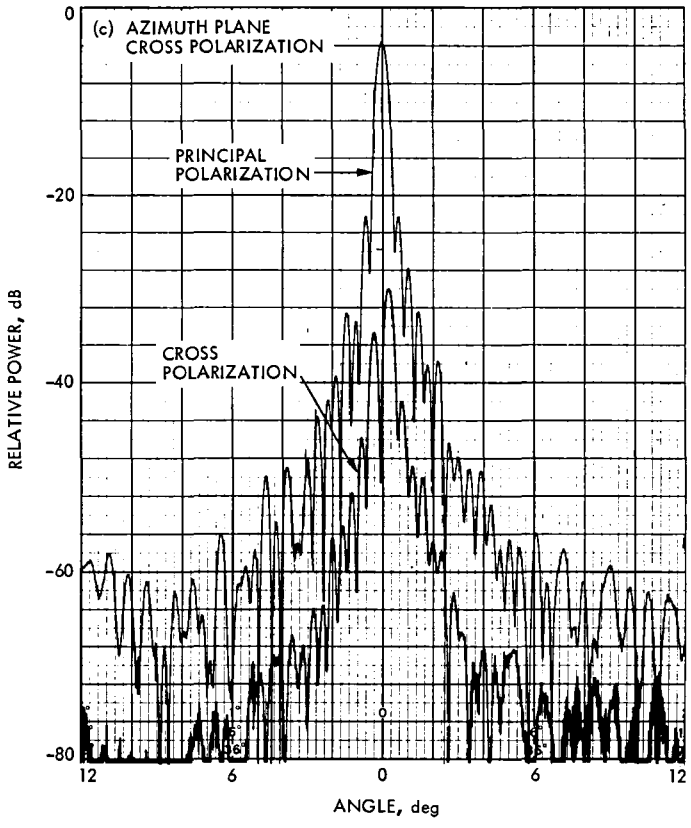


Fig. 9 (contd)

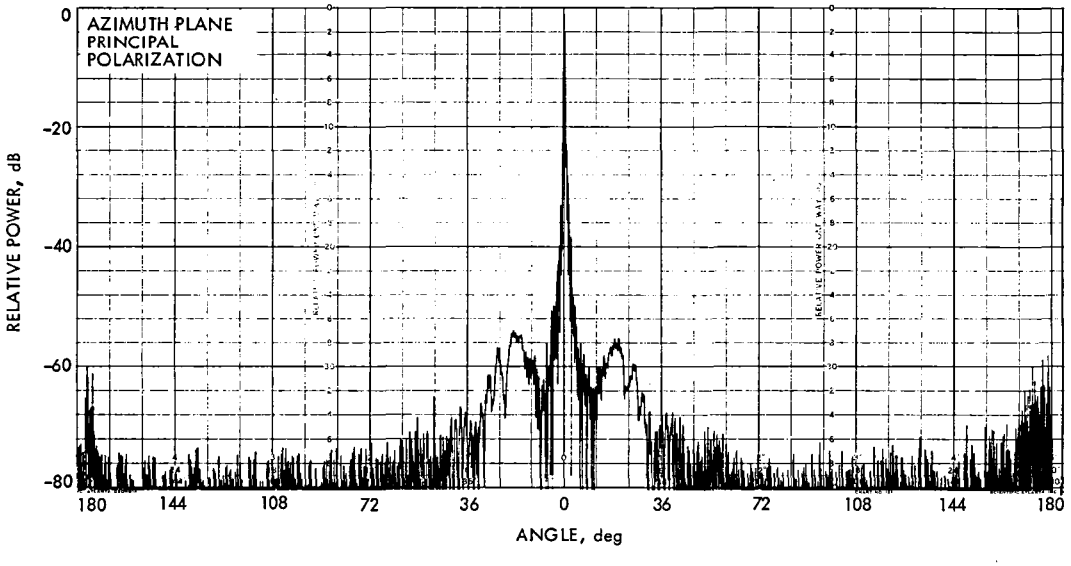


Fig. 10. Wide angle radiation characteristics

Simplified Syndrome Decoding of (n, 1) Convolutional Codes

I. S. Reed
Department of Electrical Engineering
University of Southern California

T. K. Truong
Communication System Research

This paper presents a new syndrome decoding algorithm for the (n, 1) convolutional codes (CC) that is different and simpler than the previous syndrome decoding algorithm of Schalkwijk and Vinck. The new algorithm uses the general solution of the polynomial linear Diophantine equation for the error polynomial vector E(D). This set of Diophantine solutions is a coset of the CC space. A recursive or Viterbi-like algorithm is developed to find the minimum weight error vector $\hat{E}(D)$ in this error coset. An example illustrating the new decoding algorithm is given for the binary nonsymmetric (2, 1) CC.

I. Introduction

In this paper the syndrome decoding algorithm invented in 1976 by Schalkwijk and Vinck (Ref. 1) is simplified and generalized to all $(n, 1)$ convolutional codes (CC), both systematic and nonsystematic. Indications are given also of how the techniques used here can be further extended to apply to any (n, k) CC for which a parity check polynomial matrix can be found.

Following Refs. 2 and 3, the input $X(D)$ and output sequences $Y_1(D), \cdots, Y_n(D)$ of an $(n, 1)$ CC are formal power series of finite length over a finite field $GF(q)$ in the unit delay operator D. The input $X(D)$ and the output $Y(D) = [Y_1(D), \cdots, Y_n(D)]$, as a vector, are connected by a $1 \times n$ polynomial generator matrix $G(D)$ of form

$$G(D) = [G_1(D), \cdots, G_n(D)], \tag{1}$$

where $G_k(D)$ are monic polynomials of finite degree in D over $GF(q)$. This relationship is

$$Y(D) = X(D)G(D) \tag{2}$$

The maximum degree M of the polynomials in $G(D)$ is called the memory and the constraint of the code is $L = M + 1$.

To avoid the possibility of catastrophic error propagation the important criterion of Massey and Sain (Ref. 4) is a necessity. In this paper attention is restricted to coder inverses without delay. Hence one must have

$$GCD[G_1(D), \cdots, G_n(D)] = 1$$

where GCD denotes "greatest common divisor" and $G_k(D)$ is the kth component polynomial of $G(D)$ in (1).

The parity-check matrix $H(D)$, associated with a general $k \times n$ generator matrix $G(D)$, is a $k \times (n-k)$, maximum rank matrix of polynomials over $GF(q)$ with the property

$$G(D)H^T(D) = 0 \tag{3}$$

where T denotes matrix transpose. If $G(D)$ is the generator matrix of a systematic (n, k) CC, $G(D)$ has the form $[I_k, P(D)]$ where I_k is the $k \times k$ identity matrix and $P(D)$ is a $k \times (n-k)$ matrix of polynomials in D over $GF(q)$. In this case it is easily shown (Ref. 5) that

$$H(D) = [-P^T(D), I_{n-k}] \tag{4}$$

is the appropriate parity check matrix.

The parity check matrix $H(D)$ for the general nonsystematic (n, k) CC is considerably more difficult to find. Forney in Ref. 3 develops a general procedure to find $H(D)$. However, for the $(n, 1)$ CC considered in this paper the powerful machinery of Forney is not needed. For the nonsystematic $(n, 1)$ CC with the generator matrix $G(D)$, given in (1), it is easily verified that

$$H(D) = \begin{bmatrix} G_2(D), & G_1(D), & 0 & \cdots & 0 \\ G_3(D), & 0, & G_1(D) & \cdots & 0 \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ G_n(D), & 0, & 0 & \cdots & G_1(D) \end{bmatrix} \tag{5}$$

satisfies (3) and is a parity check matrix.

Let $Z(D) = Y(D) + E(D)$ be the received code sequence in powers of the delay operator D, possibly corrupted by an error or noise sequence $E(D)$. The *syndrome* $S(D)$ of $Z(D)$ is defined by

$$S(D) = Z(D)H^T(D) \tag{6}$$

By (2) and (3) the syndrome reduces to

$$S(D) = [(Y(D) + E(D)]\,H^T(D) = E(D)H^T(D) \tag{7}$$

in terms of the noise sequence $E(D)$, only. For example by (5) and (7) the syndrome for the 1/2 rate code, (2, 1) CC, is

$$S(D) = E_1(D)G_2(D) + E_2(D)G_1(D) \tag{8}$$

For the 1/3 rate code, (3, 1) CC,

$$\begin{aligned} S(D) &= [S_1(D), S_2(D)] \\ &= [E_1(D)G_2(D) + E_2(D)G_1(D), E_1(D)G_3(D) \\ &\quad + E_3(D)G_1(D)] \end{aligned} \tag{9}$$

In Ref. 1 Schalkwijk and Vinck showed how the states of the syndrome processor of $S(D)$ in (8) could be used to form a trellis diagram for implementing a recursive algorithm similar to the Viterbi algorithm (Ref. 5). In the present paper this idea is extended in a manner which is simpler and more easily applied. The departure point for the new simplified syndrome decoding is to first find the solution of the syndrome equation (7) for the error vector $E(D)$.

This technique for finding the solution of (7) for error vector $E(D)$ is illustrated by finding solutions for $E_k(D)$ in the special cases (8) and (9). The total set in which the solutions $E_k(D)$ are to be found is the set $F[D]$ of all polynomials in operator D over $F = GF(q)$. It is well known (Ref. 6) that $F[D]$ is an integral domain (a commutative ring without zero divisors) and as a consequence satisfies many of the properties of the integers. In particular (8), (9) and more generally (7) are linear Diophantine equations over polynomials in D instead of the integers.

Using techniques precisely similar to those used for the integers, e.g., see Ref. 6, the general solution of (8) is readily found. Since $GCD[G_1(D), G_2(D)] = 1$, the Euclidean algorithm can be used to find polynomials $\alpha_1(D)$ and $\alpha_2(D)$ such that

$$\alpha_1(D)G_2(D) + \alpha_2(D)G_1(D) = 1.$$

In terms of $\alpha_k(D)$ the general solution of (8) is

$$\begin{aligned} E_1(D) &= \alpha_1(D)S(D) + G_1(D)t(D) \\ E_2(D) &= \alpha_2(D)S(D) + G_2(D)t(D) \end{aligned} \tag{10}$$

where $t(D)$ is an arbitrary polynomial in $F[D]$.

To find the Diophantine solution of (9), first eliminate $E_1(D)$ from the two components $S_k(D)$ for $(k = 1, 2)$ by multiplying the first component by $G_3(D)$ and the second by $G_2(D)$. The resulting equation after subtraction is

$$\begin{aligned} [E_2(D)G_3(D) - E_3(D)G_2(D)]\,G_1(D) &= G_3(D)S_1(D) \\ &\quad - G_2(D)S_2(D) \end{aligned} \tag{11}$$

Observe that in terms of the original computation of $S(D)$ in (6),

$$\begin{aligned} G_3(D)S_1(D) - G_2(D)S_2(D) &= G_3(D)[Z_1(D)G_2(D) \\ &\quad + Z_2(D)G_1(D)] \\ &\quad - G_2(D)[Z_1(D)G_3(D) \\ &\quad + Z_3(D)G_1(D)] \\ &= [Z_2(D)G_3(D) \\ &\quad - Z_3(D)G_2(D)]\, G_1(D) \end{aligned}$$

so that the right side of (11) is always divisible by $G_1(D)$. Dividing (11) by $G_1(D)$ yields

$$E_2(D)G_3(D) - E_3(D)G_2(D) = R(D) \qquad (12)$$

where

$$R(D) = \frac{[G_3(D)S_1(D) - G_2(D)S_2(D)]}{G_1(D)}, \qquad (13)$$

a polynomial in D over $GF(q)$. The greatest common divisor of the $G_k(D)$'s must equal one by the Massey and Sain criterion. This criterion is achieved for this case by assuming $GCD\,[G_2(D), G_3(D)] = 1$. Thus (12) has a solution similar to (10), namely,

$$\begin{aligned} E_2(D) &= \beta_3(D)R(D) + G_2(D)t(D) \\ E_3(D) &= -\beta_2(D)R(D) + G_3(D)t(D) \end{aligned} \qquad (14)$$

where $\beta_2(D)$ and $\beta_3(D)$ are a particular solution to

$$\beta_2(D)G_2(D) + \beta_3(D)G_3(D) = 1 \qquad (15)$$

and $t(D)$ is an arbitrary polynomial in D over $GF(q)$. Finally to find $E_1(D)$ substitute (14) into the components of $S(D)$ in (9) and solve for $E_1(D)$ by multiplying the first equation by $\beta_2(D)$ and the second by $\beta_3(D)$. This yields

$$E_1(D) = S_1(D)\beta_2(D) + S_2(D)\beta_3(D) - G_1(D)t(D) \qquad (16)$$

where $\beta_2(D)$ and $\beta_3(D)$ satisfies (15). Equations (14) and (16) with (15) where $t(D) \in F[D]$ constitute the general solution of (7) for $E(D)$ of the (3, 1) CC. The above Diophantine techniques extend to yield solutions to all $(n, 1)$ CC. In fact it is readily demonstrated that the general solution of (7) for $E(D)$ is the linear function

$$E(D) = L\,[t(D)] = L_0(D) + L_1(D)t(D). \qquad (17)$$

for all $t(D)$ in $F[D]$. The set of all $L\,[t(D)]$ is a coset of the $(n, 1)$ CC code space $\{L_1(D)t(D) | t(D) \in F[D]\}$.

In order to use syndrome decoding to find a maximum likelihood estimate (MLE) $\hat{E}(D)$ of the actual error sequence, both the weight of the sequence and the channel need to be defined. For an $(n, 1)$ CC a possible error sequence is of form $E(D) = [E_1(D), E_2(D), \cdots, E_n(D)]$ where $E_k(D)$ are finite degree polynomials over $GF(q)$. The Hamming weight of $E(D)$ is

$$W_H[E(D)] = \sum_{k=1}^{n} W_H[E_k(D)]$$

where $W_H[E_k(D)]$, the Hamming of polynomial $E_k(D)$, is the number of nonzero coefficients of $E_k(D)$. Assume the channel over which $Y(D)$ is sent is approximated by a q-ary channel (see Ref. 2, Sec. 7.2).

If $\deg[X(D)] \leqslant L - 1$, the codeword $Y(D) = [Y_1(D), \cdots, Y_n(D)]$ is the Lth truncation of an $(n, 1)$ CC (Ref. 2, p. 203). Each component $Y_k(D)$ has degree $\leqslant M + L - 1$ where M is memory of code. For a truncated $(n, 1)$ CC transmitted over a q-ary symmetric channel it is evident that the MLE of an error vector is $\hat{E}(D)$ such that

$$W_H(\hat{E}) = \underset{t(D)}{\text{Min}}\ \{L\,[t(D)]\} \qquad (18)$$

where $L\,[t(D)]$ is the linear Diophantine solution (17) for $E(D)$ of syndrome equation (7)

The above procedure for finding the MLE $\hat{E}(D)$ is equivalent to the standard syndrome decoding technique used for block codes, e.g., see Ref. 7. In the next section a recursive or Viterbi-like algorithm is developed to efficiently find $\hat{E}(D)$, the estimate of the error sequence.

II. Recursive Syndrome Decoding

The new technique of recursive syndrome decoding is presented by example, with the same nonsystematic (2, 1) CC used in Ref. 1. For this code,

$$G(D) = [G_1(D), G_2(D)] = [1 + D^2, 1 + D + D^2]$$

and

$$H(D) = [G_2(D), G_1(D)] = [1 + D + D^2, 1 + D^2]$$

are the generating and parity-check matrices, respectively. By (8) the syndrome is

$$\begin{aligned} S(D) &= Z_1(D)(1 + D + D^2) + Z_2(D)(1 + D^2) \\ &= Z_1(D) + Z_1^{(1)}(D) + Z_1^{(2)}(D) + Z_2(D) + Z_2^{(2)}(D) \end{aligned} \tag{19}$$

Note that terms such as $Z_1^{(1)}(D)$ in this expression can be regarded either as a delayed version of $Z_1(D)$ or a right shift of sequence $Z_1(D)$ when viewed as a sequence proceeding from left to right. Finally the Diophantine solutions (10) of the syndrome equation (8) for error sequences $E_1(D)$ and $E_2(D)$ are

$$\begin{aligned} E_1(D) &= DS(D) + (1 + D^2)t(D) \\ &= S^{(1)}(D) + t(D) + t^{(2)}(D) \end{aligned}$$

and

$$\begin{aligned} E_2(D) &= (1 + D)S(D) + (1 + D + D^2)t(D) \\ &= S(D) + S^{(1)}(D) + t(D) + t^{(1)}(D) + t^{(2)}(D) \end{aligned} \tag{20}$$

since

$$\alpha_1(D) = D \quad \text{and} \quad \alpha_2(D) = 1 + D$$

constitute a particular solution of

$$\alpha_1(D)G_1(D) + \alpha_2(D)G_2(D) = 1.$$

Schalkwijk and Vinck (Ref. 1) used the states of the sequential circuit, used to form the syndrome $S(D)$, for the states of their trellis diagram. Here the states of the trellis diagram are obtained from the states of the shift register needed in (20) to obtain $t^{(1)}(D)$ and $t^{(2)}(D)$ from $t(D)$. A block diagram of a shift register to produce the delayed versions, $t^{(1)}(D)$ and $t^{(2)}(D)$, from $t(D)$ is shown in Fig. 1. The state table of this shift register, when regarded as a sequential circuit, is given in Fig. 2. Finally in Fig. 3 the trellis diagram, associated with this state table is presented. A solid-line transition in Fig. 3 corresponds to the input $t(D) = 0$; a dashed-line transition corresponds to the input $t(D) = 1$.

The new Viterbi-like syndrome decoding algorithm is illustrated by example in Fig. 4. For this example assume that the message to be transmitted is the six-bit message

$$X(D) = [0\ 1\ 0\ 0\ 1\ 0]$$

so that the truncation length is $L = 6$. By (2)

$$Y_1(D) = [0\ 1\ 0\ 1\ 1\ 0\ 1\ 0]$$

and

$$Y_2(D) = [0\ 1\ 1\ 1\ 1\ 1\ 1\ 0]$$

are the two components of $Y(D)$. Thus the scalar representation of the code word is

$$C = [0\ 0\ 1\ 1\ 0\ 1\ 1\ 1\ 1\ 1\ 0\ 1\ 1\ 1\ 0\ 0]$$

of overall code length $n(L + M) = 2(6 + 2) = 16$ (see Ref. 2, pp. 201–203).

Let the received codeward be

$$R = [1\ 0\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 0\ 0]$$

Then

$$Z_1(D) = [1\ 1\ 0\ 1\ 1\ 1\ 1\ 0]$$

and

$$Z_2(D) = [0\ 1\ 0\ 1\ 1\ 1\ 1\ 0]$$

are the two components of received message $Z(D)$. Using (19) this syndrome sequence is

$$\begin{aligned} S(D) &= 1 + D + D^4 + D^5 + D^6 + D^7 \\ &= [1\ 1\ 0\ 0\ 1\ 1\ 1\ 1\ 0] \end{aligned}$$

both as a polynomial in D over $GF(2)$ and as a simple finite sequence. The latter representation of $S(D)$ is shown in Fig. 4 with its digits placed over the transition paths of the trellis.

Let us proceed briefly through the trellis. The syndrome is set to 0 prior to stage 0 so that at stage 0, $S^{(1)}(D) = 0$, and

$S(D) = 1$. Likewise it is assumed initially that the shift register in Fig. 1 is cleared so that $t^{(1)}(D) = t^{(2)}(D) = 0$. This puts the algorithm at stage 0 at stage $a = [0, 0]$. The label on branch $t(D) = 0$ is found by substituting these values in (20). That is,

$$E_1(D) = S^{(1)}(D) + t(D) + t^{(2)}(D) = 0 + 0 + 0 = 0$$

$$E_2(D) = S(D) + S^{(1)}(D) + t(D) + t^{(1)}(D) + t^{(2)}(D)$$

$$= 1 + 0 + 0 + 0 + 0 = 1$$

Hence the label on branch $t(D) = 0$ at stage zero in $[E_1(D)$ $E_2(D)] = [0, 1]$. By the same substitution but with $t(D) = 1$, the label on the alternate branch is $[E_1(D), E_2(D)] = [1, 0]$, the componentwise complement of the previous branch. The Hamming weight of $[E_1(D), E_2(D)]$ for both these branches is 1. Note that this weight of 1 is placed immediately above states a and c at stage 1.

To illustrate the Viterbi or dynamic-programming technique for computing survivors in the trellis suppose that the algorithm is either at state a or state b at stage 2. Note from Fig. 3 that there are only two ways to state a, by a transition from b to a or a transition from a to a. At state a or b at stage 2, $S(D) = 0$ and $S^{(1)}(D) = 1$. At stage 2 at state a, $t^{(1)}(D) = 0$ and $t^{(2)}(D) = 0$ so that at branch $t(D) = 0$,

$$E_1(D) = S^{(1)}(D) + t(D) + t^{(2)}(D) = 1 + 0 + 0 = 1$$

$$E_2(D) = S(D) + S^{(1)}(D) + t(D) + t^{(1)}(D) + t^{(2)}(D)$$

$$= 0 + 1 + 0 + 0 + 0 = 1$$

The total weight of $[E_1(D), E_2(D)]$ for the minimum weight path, going through state a at stage 2, is thus $2 + 2 = 4$. However, at stage 2 at state b, $t^{(1)}(D) = 0$ and $t^{(2)}(D) = 1$, so that at branch $t(D) = 0$,

$$E_1(D) = S^{(1)}(D) + t(D) + t^{(2)}(D) = 1 + 0 + 1 = 0$$

$$E_2(D) = S(D) + S^{(1)}(D) + t(D) + t^{(1)}(D) + t^{(2)}(D)$$

$$= 0 + 1 + 0 + 0 + 1 = 0$$

Thus the total weight of $[E_1(D), E_2(D)]$ for the minimum weight path, going through state b at stage 2 is thus $3 + 0 = 3$. Since this weight is smaller than the previous weight, only the path going through state b to a at stage 2 survives. The segment of path from state a to a is deleted as shown in Fig. 4. Similarly in Fig. 4 some paths lead to equal weights or a "tie." In such a case either segment can be chosen as part of the survivor path.

The entire trellis diagram shown in Fig. 4 is completed by the rules illustrated above. At stage 9 the minimum weight path in the trellis diagram of Fig. 4 is $a\ c\ d\ b\ c\ b\ c\ d\ b\ a$. The branches of this path yield

$$\hat{E}(D) = [1\ 0\ 0\ 0\ 0\ 1\ 0\ 0, 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0]$$

$$= [1 + D^5, D^2] = [\hat{E}_1(D), \hat{E}_2(D)]$$

as the estimate of the error vector. Subtracting these estimates of the error from $Z(D)$ produces

$$\hat{Y}_1(D) = [0\ 1\ 0\ 1\ 1\ 0\ 1\ 0]$$

and

$$\hat{Y}_2(D) = [0\ 1\ 1\ 1\ 1\ 1\ 1\ 0]$$

as estimates of the transmitted coded message. Finally the inverse linear sequential circuit of Massey and Sain (Ref. 4), with equation

$$\hat{X}(D) = (1 + D)\hat{Y}_1(D) + D\hat{Y}_2(D)$$

is used to find $\hat{X}(D) = [0\ 1\ 0\ 0\ 1\ 0]$ as an estimate of the original message.

For the above example, this new syndrome decoding algorithm yields the original message. However, if the number of errors exceeds the capability of the code, at the end of the decoding period there may exist two or more paths with the same minimum error weight. In such a circumstance a decoding failure and an erasure should be declared.

III. Conclusions

In this paper a new simplified syndrome decoding algorithm for $(n, 1)$ CC is developed which utilizes the general Diophantine solution for the error vector $E(D)$ in the syndrome equation. The least weight error vector $\hat{E}(D)$ is found by a recursive Viterbi-like algorithm, similar to an algorithm conceived previously by Schalkwijk and Vinck (Ref. 1)

This new syndrome decoder appears to be comparable in complexity to the Viterbi decoder except that in the new decoder fewer comparisons are required and the control logic is considerably simpler. Another possible advantage of the new algorithm is its ability to detect decoding failures more readily than the classical Viterbi approach. A more detailed comparison of these decoders is a topic for further study.

Acknowledgment

The authors wish to thank Charles Wang of JPL for his helpful suggestions made during the preparation of this paper.

References

1. Schalkwijk, J. P. M., and Vinck, A. J., "Syndrome decoding of binary rate–1/2 convolutional codes," *IEEE Trans. Comm.*, COM-24, pp. 977–985, 1976.
2. McEliece, R. J., *The Theory of Information and Coding*, Addison-Wesley Publishing Co., Reading, Mass., 1974.
3. Forney, G. D., "Convolutional Codes I: algebraic structure," *IEEE Trans. Inform.*, IT-16, pp. 720–738, 1970.
4. Massey, J. L., and Sain, M. K., "Inverses of linear circuits," *IEEE Trans. Comput.*, C-17, pp. 330–337, 1968.
5. Viterbi, A., and Omura, J., *Digital Communication and Coding*, McGraw-Hill Book Co., New York, 1978.
6. Nagell, T., *Introduction to Number Theory*, John Wiley and Sons, New York, 1951.
7. Peterson, W. W., and Weldon, E. J., Jr., *Error-Correcting Codes*, 2nd ed., M.I.T. Press, Cambridge, Mass., 1972.

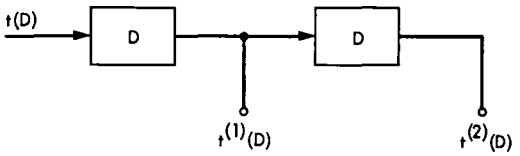


Fig. 1. Shift register to generate delayed versions of *t(D)*

$t^{(1)}(D), t^{(2)}(D)$		t(D) = 0	t(D) = 1
a = 0	0	0 0	1 0
b = 0	1	0 0	1 0
c = 0	0	0 1	1 1
d = 1	1	0 1	1 1

Fig. 2. State table of shift register for *t(D)*

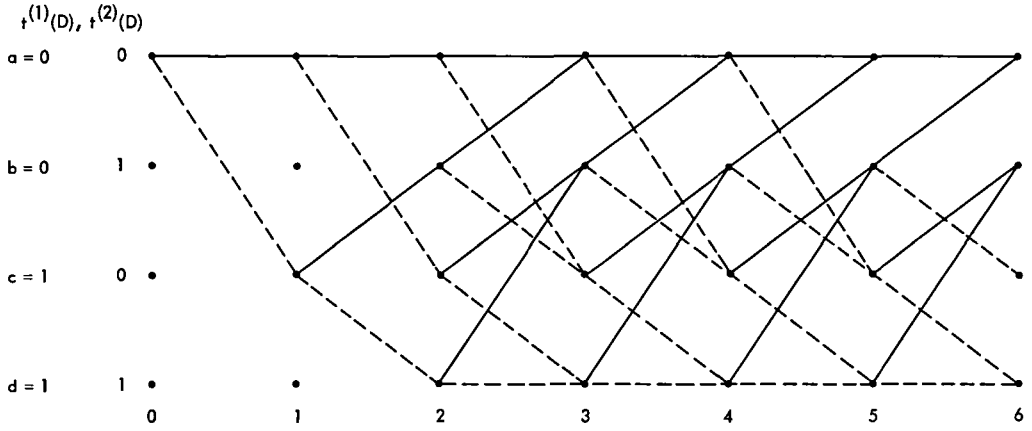


Fig. 3. Trellis diagram of shift register for *t(D)*. Input *t(D)* = 0 is represented by a solid line. *t(D)* = 1 is represented by a dashed line

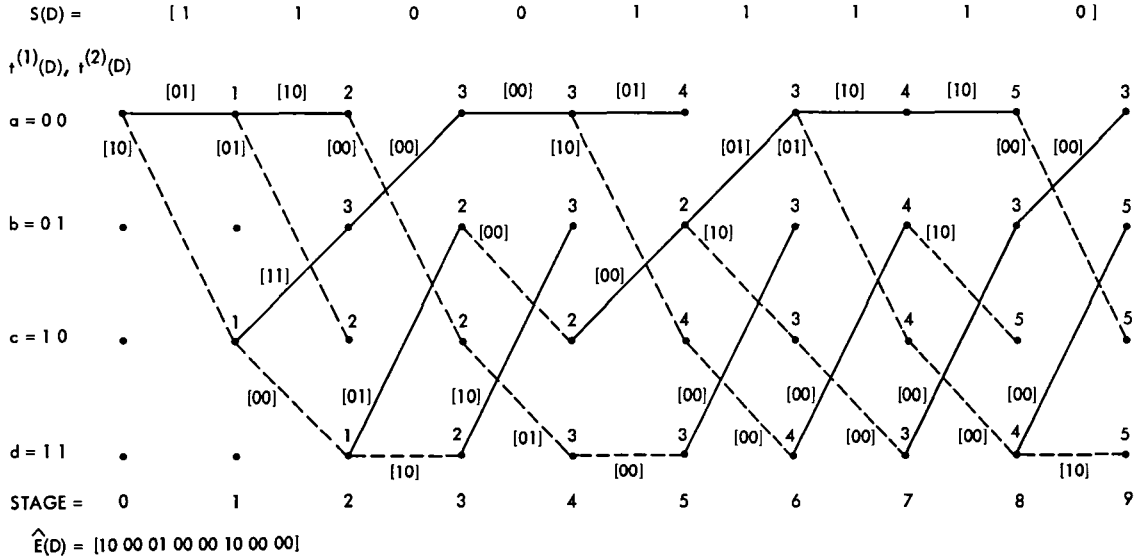


Fig. 4. New Viterbi-like syndrome decoding algorithm. Each branch of trellis is labeled with $[E_1(D), E_2(D)]$ where $E_1(D) = S^{(1)}(D)+t(D)+t^{(2)}(D)$ and $E_2(D) = S(D)+S^{(1)}(D)+t^{(1)}(D)+t^{(2)}(D)$. Each node at *k* is labeled with $W_H[E_1(D), E_2(D)]$ up to time k

Node Synchronization for the Viterbi Decoder

G. Lorden[1], R. McEliece[2], and L. Swanson
Communications Systems Research Section

At very low signal-to-noise ratios such as those that Voyager 2 will encounter at Uranus, the performance of the Reed-Solomon/Viterbi concatenated coding system could be seriously degraded by loss of node synchronization by the Viterbi decoder. In this article we will describe this problem and show that it can be almost completely avoided with a simple outboard hardware "node synchronizer." This device makes statistical decisions about node sync based on the hard-quantized undecoded data stream. We show that in a worst-case Voyager-like environment, our method will detect and correct a true loss of node sync (thought to be a very rate event) within several hundred bits; many of these outages will be corrected by the RS code. At the same time, the mean time between false alarms for our technique is on the order of several years.

I. Introduction

This paper deals with the problem of detecting and correcting losses of node synchronization in convolutionally encoded data. We are motivated by our desire to restore the loss of telemetry on future NASA deep-space missions which has been predicted in Refs. 1 and 2, and has been seen in hardware tests (Ref. 3). For definiteness, this report will only deal with the problem as it occurs on the present Voyager 2 mission, but our results will apply equally to any mission in which the telemetry is protected by a Reed-Solomon/Viterbi concatenated coding system.

On Voyager, the high-rate downlink telemetry is protected by a $K = 7$, rate 1/2 convolutional code concatenated with a depth-4 interleaved (255,223) Reed-Solomon code. In principle this combination provides excellent error protection (bit error probability 1.0E-6) for Voyager's highly sensitive imaging data, at bit signal-to-noise ratios as low as 2.9 dB. Since the rate of the outer code is 223/255 = -0.6 dB, when the overall SNR is 2.9 dB, the inner convolutional code is operating at about 2.3 dB.

However, in practice, the performance of the concatenated system is significantly worse than theoretical predictions. One problem is carrier-loop jitter, which degrades performance by 0.5 dB or more (Ref. 1). This means that if the system bit SNR remains at the nominal 2.9-dB value, the inner convolutional code must operate at less than 2.0 dB. This is a value much lower than that for which the DSN's hardware Viterbi decoders were designed. In this demanding environment, the Viterbi decoder's internal node synchronization hardware, whose function is to detect and correct true external losses of node sync, is prone to produce flase alarms i.e., spurious losses of node sync, and send useless data to the Reed-Solomon decoder until node sync is reestablished. In Ref. 2 it was shown that this hardware problem can degrade Voyager's performance by a further 1.0 dB or more.

[1]Consultant, Caltech Mathematics Department
[2]Consultant, Caltech Electrical Engineering Department

In this article we shall show that this data loss due to spurious node sync loss in the Viterbi decoder is completely avoidable. Our proposed solution involves disabling the Viterbi decoder's internal synchronization hardware and implementing an external node sync algorithm. Our algorithm is easy to implement and depends on likelihood calculations based on observations of the hard-quantized encoded data stream. In a worst-case Voyager environment, our method will detect and correct a true loss of node sync within several hundred bits; many of these outages will be corrected by the Reed-Solomon code. On the other hand, the mean time between false alarms for our technique (which is independent of the SNR) is on the order of several years. Thus for practical purposes our technique introduces no false alarms, and the system SNR loss due to node sync problems will be eliminated, with no loss of protection against true node sync losses. As an outboard hardware device, our algorithm could be implemented on a single DSN standard single-board computer such as the iSBC 86/12, at least at data rates up to 20 kbps.

The paper is divided into three sections. In Section II, we present a functional description of our algorithm, together with a summary of the relevant mathematics. In Section III, we present some numerical performance results for our technique. They will quantify the assertions made above (mean time between false alarms, probability of uncorrectable errors due to true loss of sync, etc.). We also include in the appendices some background information.

II. The Up-Down Counter

We adopt the following model, which has been found to be very accurate for coherent deep-space communication (Ref. 4). The information to be transmitted via the convolutional code (which in Voyager is already encoded) is a sequence $\ldots M_{-1}, M_{-2}, M_0, M_2, M_4, \ldots$ of independent identically distributed random variables, each equally likely to be 0 or 1. The encoded stream $\ldots, C_{-3}, C_{-2}, C_{-1}, C_0, C_1, \ldots$ is defined by the encoding equations[3]

$$C_{2k} = M_{2k} + M_{2k-2} + M_{2k-4} + M_{2k-6} + M_{2k-12} \pmod 2 \quad (1)$$

$$C_{2k-1} = M_{2k} + M_{2k-4} + M_{2k-6} + M_{2k-10} + M_{2k-12} \pmod 2 \quad (2)$$

[3]We shall illustrate all of our results for the NASA standard $K = 7$, rate 1/2 convolutional code, but everything generalizes easily to any rate 1/2 convolutional code.

We also define the ± versions of the encoded stream:

$$D_j = \begin{cases} +1 \text{ if } C_j = 0 \\ -1 \text{ if } C_j = 1. \end{cases}$$

The encoded bits $\{D_k\}$ are used to modulate a radio frequency signal, which is transmitted by Voyager to Earth. After detection and demodulation, a sequence $\{\hat{D}_k\}$ is received, where $\hat{D}_k = D_k + Z_k$. The sequence $\{Z_k\}$ is the *error* sequence. If the noise process is additive white Gaussian noise, then the sequence $\{Z_k\}$ is i.i.d., the common distribution being Gaussian, mean zero and variance $\sigma^2 = 1/\mu$, where μ is twice the symbol SNR.

The Viterbi decoder attempts to recover the message bits from the noisy code sequence $\{\hat{D}_k\}$. It does this by making a very efficient maximum likelihood estimate of each of the message bits $\{M_{2j}\}$. However, in order to operate, the Viterbi decoder *must have node sync*, i.e., it must know which of the received symbols have *even* subscripts and which have *odd* subscripts. Of course there are only two possibilities, but if the wrong hypothesis is made, the output of the Viterbi decoder will bear no useful relationship to the message stream $\{M_{2j}\}$.

Our algorithm will provide node sync information for the Viterbi decoder. It is based on the *hard-quantized* received sequence $\{R_k\}$, where

$$R_k = \begin{cases} 0 \text{ if } \hat{D}_k \text{ is} \geqslant 0 \\ 1 \text{ if } \hat{D}_k \text{ is} < 0. \end{cases}$$

Clearly $R_k = C_k + E_k \pmod 2$, where $\{E_k\}$ is i.i.d, and $E_k = 0$ or 1. The *error probability* $P_e = Pr\{E_r = 1\}$ is given by

$$P_e = \frac{1}{\sqrt{2\pi}} \int_{\mu}^{\infty} e^{-t^2/2}\, dt, \quad (3)$$

where as before $\mu = 2E_S/N_0$.

Associated with the hard-quantized received sequence $\{R_k\}$ is the *syndrome*, or *parity-check sequence* $\{X_k\}$ (Ref. 5):

$$X_k = R_k + R_{k-1} + R_{k-3} + R_{k-4} + R_{k-5} + R_{k-6} + R_{k-7} + R_{k-10} + R_{k-12} + R_{k-13}. \quad (4)$$

If there are no errors, i.e., if $E_k = 0$ for all k, then it follows from (1) and (2) that $X_k = 0$ for all even subscripts k. In fact we have explicitly

$$X_k = E_k + E_{k-1} + E_{k-3} + E_{k-4} + E_{k-5} + E_{k-6} + E_{k-7} + E_{k-10} + E_{k-12} + E_{k-13} \quad k \text{ even.} \tag{5}$$

For odd k, (1) and (2) show that the X_k's are independent and take the values 0 and 1 with probability 1/2 each, regardless of the $\{E_k\}$ sequence. (This is shown in detail in Appendix A.)

For even k, (5) shows that E_k will be zero if and only if an even number of $\{E_k, E_{k-1}, E_{k-3}, E_{k-4}, E_{k-5}, E_{k-6}, E_{k-7}, E_{k-10}, E_{k-12}, E_{k-13}\}$ are one. The probability of this is easily seen to be

$$\pi = \frac{1 + (1 - 2p_e)^{10}}{2}. \tag{6}$$

The foregoing describes the distribution of the received sequence and of the parity-check sequence in case node synchronization is maintained. This will be called the *in-sync hypothesis*. The *out-of-sync hypothesis* describes the situation when node synchronization is in error. In this case the R_i's and X_i's behave as though the subscripts were shifted by one. Thus, under the out-of-sync hypothesis it is the odd parity checks that are correct with probability π and the even ones that are purely random.

We assume that node synchronization has been acquired and maintained and that the in-sync hypothesis is initially true. We wish to monitor the received sequence so as to detect loss of sync, i.e., a sudden change making the received sequence obey the out-of-sync hypothesis. A method for doing this simply can be based on a general statistical technique (Ref. 6) for detecting a change in distribution. To apply this technique it is necessary to simplify the model by assuming that the parity checks X_i are independent. They are not independent for even i, but the dependence between widely separated X_i's is slight, so calculations based on this assumption should be illustrative.

The method for detecting loss of sync is based on a *counter* with increments

$$L(X_n) = \log \frac{q(X_n)}{p(X_n)},$$

where $p(X_n)$ and $q(X_n)$ are the likelihoods of X_n under the in-sync and out-of-sync hypotheses, respectively.

As shown above

$$p(X_n) = \begin{cases} 1/2 \text{ if } n \text{ is odd} \\ \pi \text{ if } n \text{ is even and } X_n = 0 \\ 1 - \pi \text{ if } n \text{ is even and } X_n = 1 \end{cases}$$

whereas $q(X_n)$ reverses the odd and even cases. Thus,

$$L(X_n) = \begin{cases} (-1)^{n+1} \log(2\pi) & \text{if } X_n = 0 \\ (-1)^{n+1} \log(2(2 - \pi)) & \text{if } X_n = 1. \end{cases}$$

The counter is defined by

$$T_0 = 0$$

$$T_n = \max([T_{n-1} + L(X_n)], 0), n \geq 1.$$

A threshold $\gamma > 0$ is chosen, and the process stops the first time $T_n \geq \gamma$. Since at this point there is a substantial likelihood ratio in favor of the out-of-sync hypothesis, the inference is made that loss of node synchronization has occurred. This loss of node sync can be remedied by either adding or deleting a channel symbol; the node synchronizer will alternate adding and deleting symbols in order to avoid ruining frame boundaries.

The performance of such a counter for a particular γ is characterized by two *average run lengths* (ARL's).

(1) The *short* ARL. This is the average number of pairs of symbols needed to reach the threshold if the out-of-sync hypothesis is true from the beginning, and

(2) The *long* ARL. This is the average number of pairs of symbols needed to reach the threshold if the in-sync hypothesis remains true.

The *short* ARL gives an upper bound on the average time between loss of sync and its detection, since whenever loss of sync occurs the counter has a nonnegative value. The *long* ARL (or its reciprocal) describes the frequency of false detection of loss of node synchronization.

An exact determination of the ARL's will be made in Section III, for a very slight modification of the scheme described. It is instructive, however, particularly for later comparisons, to consider their asymptotic behavior as $\gamma \to \infty$. It turns out that

$$\text{short ARL} \sim \frac{\gamma}{I}$$

and

$$\text{long ARL} \sim C e^{\gamma},$$

so that

$$\text{short ARL} \sim \frac{\log(\text{long ARL})}{I} . \qquad (7)$$

Here the constant I, which is much more critical than C, is the Kullback-Leibler *information number* per pair of symbols. This number is simply the average increment of the counter per pair of symbols when the out-of-sync hypothesis is true. Since each pair of symbols generates one even X_n and one odd X_n,

$$I = I_1 + I_2 ,$$

where

$$I_1 = q(0|n \text{ odd}) \log 2\pi + q(1|n \text{ odd}) \log(2(1-\pi))$$

$$= \pi \log(2\pi) + (1-\pi) \log(2(1-\pi)),$$

and

$$I_2 = q(0|n \text{ even})(-\log 2\pi) + q(1|n \text{ even})(-\log(2(1-\pi))$$

$$= -1/2 \log(4\pi(1-\pi)) .$$

Table 1 shows the dependence of the information numbers on P_e, the symbol error probability. It should be noted that the information numbers decrease substantially as the symbol error probability becomes larger. Thus, the ARL's become less favorable as p_e increases.

If one makes the very slight change of inspecting the counter only at even n, i.e., once for each pair of symbols, then a simpler description of the counter is possible. This is because two consecutive X_n's = 0 (successful parity checks) yield a net change of zero, as do consecutive X_n's = 1. If, however, one of the pair of X_n's is 0 and the other 1, then the total increment of the counter is easily seen to be $\pm \log \pi/(1-\pi)$, with + if and only if the odd $X_n = 0$ (i.e., both a successful check for the out-of-sync hypothesis and a failure for the in-sync hypothesis). Thus the counter moves up and down by a fixed step size and standard random walk formulas (Ref. 7, p. 351) can be used to derive exact formulas for the ARL's under the simplifying assumption of independence. Assuming without loss of generality that

$$\gamma = m \log \frac{\pi}{1-\pi},$$

m an integer, one has

$$\text{short ARL} = (\pi - 0.5)^{-1} \left(m + \frac{1 - [\pi/(1-\pi)]^{-m}}{1 - [\pi/(1-\pi)]} \right)$$

and

$$\text{long ARL} = (\pi - 0.5)^{-1} \left(\frac{[\pi/(1-\pi)]^{m} - 1}{1 - (1-\pi)/\pi} - m \right)$$

Table 2 illustrates the relationship between the two ARL's as a function of m for two symbol error probabilities.

A. Counter with Memory

A simple parity check counter does a fairly good job of node synchronization in case of high SNR (Ref. 5). But in the case of high symbol error probability, the probability of parity check error is quite high. For example, symbol error probability 0.1 corresponds to $\pi = 0.45$. Thus, in-sync data with $p_e = 0.1$ will fail an even parity check with probability 0.45, while out-of sync data will fail the even parity checks with probability 0.5. A counter to distinguish between distributions which are so close will either require a long time to react to incorrect sync or have a large probability of incorrect change. Performance numbers for such a counter are indicated in Table 2.

If the parity checks were independent, there would be no way to improve this performance (Ref. 6). But parity checks are not independent. This is because, under the in-sync hypothesis, one channel symbol error changes the value of five even parity checks. Under the in-sync hypothesis, for example, an isolated error in the (n-12)th channel symbol will cause parity check failures at time n, $n-2$, $n-6$, $n-8$, and $n-12$. Thus a long sequence of successful parity checks would lead us to believe that another success is on the way, while the sequence $X_{n-12} = 1, X_{n-10} = 0, X_{n-8} = 1, X_{n-6} = 1, X_{n-4} = 0, X_{n-2} = 1$ would lead one to believe that X_n is very likely to be 1. The reason a log-likelihood counter works so well in the independent case is that by adding logarithms we multiply their arguments. After receiving parity checks $X_1 = x_1$, $X_2 = x_2, \ldots, X_n = x_n$, where each x_i is zero or one, the counter contains

$$\log \frac{q(x_1)}{p(x_1)} + \log \frac{q(x_2)}{p(x_2)} + \cdots + \log \frac{q(x_n)}{p(x_n)} = \log \frac{\prod_{i=1}^{n} q(x_i)}{\prod_{i=1}^{n} p(x_i)} .$$

If the parity checks were independent, this would be exactly

$$\log\left(\frac{q(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n)}{p(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n)}\right),$$

the log of the ratio of the likelihood of the string

$$(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n)$$

under the two hypotheses, which is statistically optimal for detecting loss of synchronization.

In the case of noisy convolutionally encoded data, the probability of a string of parity checks is not just the product of the probabilities of the individual parity checks. Updating the p probability of a string requires the conditional probability $p(X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \cdots, X_1 = x_1)$. Therefore, a counter with increment

$$\log\left(\frac{q(X_n = x_n | X_{n-1} = x_{n-1}, \cdots, X_1 = x_1)}{p(X_n = x_n | X_{n-1} = x_{n-1}, \cdots, X_1 = x_1)}\right)$$

would, after step r, contain

$$\sum_{n=1}^{r} \log \frac{q(X_n = x_n | X_{n-1} = x_{n-1}, \cdots, X_1 = x_1)}{p(X_n = x_n | X_{n-1} = x_{n-1}, \cdots, X_1 = x_1)} =$$

$$\log \frac{\prod_{n=1}^{r} q(X_n = x_n | X_{n-1} = x_{n-1}, \cdots, X_1 = x_1)}{\prod_{n=1}^{r} p(X_n = x_n | X_{n-1} = x_{n-1}, \cdots, X_1 = x_1)} =$$

$$\log \frac{q(X_r = x_r, X_{r-1} = x_{r-1}, \cdots, X_1 = x_1)}{p(X_r = x_r, X_{r-1} = x_{r-1}, \cdots, X_1 = x_1)},$$

exactly the log of the quantity we desire. Of course, a real counter won't take into account a past of indefinite length. But a (possibly large) integer m can be chosen, and a counter constructed whose increment at time n is

$$\log\left(\frac{q(X_n = x_n | X_{n-1} = x_{n-1}, \cdots, X_{n-m} = x_{n-m})}{p(X_n = x_n | X_{n-1} = x_{n-1}, \cdots, X_{n-m} = x_{n-m})}\right) \tag{8}$$

As might be expected, for large m, the information number approaches the information number for the hypothetical counter based on the indefinite past of the parity check sequence. This is verified in Appendix A. Moreover, the information numbers obtained using the parity check sequence are identical with information numbers obtainable from the hard quantized received sequence. This means that there is no loss of information or efficiency in using the parity check sequence instead of the hard quantized received sequence to detect loss of node synchronization. This is also shown in Appendix A. Also, Appendix A shows that the counter increments depend only on parity checks of the same (odd or even) type, i.e., (8) is unchanged if the given $(X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \cdots)$ is replaced by $(X_{n-2} = x_{n-2}, X_{n-4} = x_{n-4}, \cdots)$.

We will describe the counter in terms of the number of parity checks used to determine the counter increment, and we will call this k. For example, the "simple" (memoryless) parity check counter corresponds to $k = 1$, while the counter for $k = 8$ has counter increments

$$\log\left(\frac{q(X_n = x_n | X_{n-2} = x_{n-2}, \cdots, X_{n-14} = x_{n-14})}{p(X_n = x_n | X_{n-2} = x_{n-2}, \cdots, X_{n-14} = x_{n-14})}\right)$$

In general, the counter using k parity checks has increment

$$\log\left(\frac{q(X_n = x_n | X_{n-2} = x_{n-2}, \cdots, X_{n-2k+2} = x_{n-2k+2})}{p(X_n = x_n | X_{n-2} = x_{n-2}, \cdots, X_{n-2k+2} = x_{n-2k+2})}\right)$$

The system which we have investigated in detail is therefore a system which takes hard-quantized received channel symbols $R_1, R_2, \cdots$, creates parity checks $X_{14}, X_{15}, \cdots$, with

$$X_i = R_i + R_{i-1} + R_{i-3} + R_{i-4} + R_{i-5} + R_{i-6}$$

$$+ R_{i-7} + R_{i-10} + R_{i-12} + R_{i-13},$$

and keeps a counter whose increment at time n is

$$\log\left(\frac{q(X_n = x_n | X_{n-2} = x_{n-2}, X_{n-4} = x_{n-4}, \cdots, X_{n-2k+2} = x_{n-2k+2})}{p(X_n = x_n | X_{n-2} = x_{n-2}, X_{n-4} = x_{n-4}, \cdots, X_{n-2k+2} = x_{n-2k+2})}\right).$$

In order to design this counter, we need to know

$$p(X_n = x_n | X_{n-2} = x_{n-2}, \cdots, X_{n-2k+2} = x_{n-2k+2})$$

and q of the same event for even and odd n. To calculate these values, remember that for in-sync data, the even parity checks depend only on the sequence of channel symbols errors. So given a probability p_e of symbols error, we can calculate the probability that

$$E_n = e_n, E_{n-1} = e_{n-1}, \cdots, E_{n-2k-11} = e_{n-2k-11}$$

for each sequence e_n, e_{n-1}, $\cdots$, $e_{n-2k-11}$, and use these probabilities to calculate

$$p(X_n = x_n, \cdots, X_{n-2k+2} = x_{n-2k+2})$$

and

$$p(X_{n-2} = x_{n-2}, \cdots, X_{n-2k+2} = x_{n-2k+2})$$

and find

$$p(X_n = x_n | X_{n-2} = x_{n-2}, \cdots, X_{n-2k+2} = x_{n-2k+2})$$

for even n.

For odd n, just as in the case of the simple counter, the X_n's are independent with $p(X_n = 0) = 1/2$. Thus

$$p(X_n = x_n | X_{n-2} = x_{n-2}, \cdots, X_{n-2k+2} = x_{n-2k+2}) = 1/2$$

for odd n.

To calculate the probabilities q associated with the out-of-sync hypothesis, just exchange even with odd in the above argument. From these values, we can calculate the counter increments. We did this, assuming a channel symbol error rate of 0.1, which corresponds to a Viterbi decoded bit error rate of 5×10^{-3}, the standard for imaging data. Voyager's data rate has been adjusted so that this is the largest channel symbol error rate which will be encountered.

B. Information Numbers and Run Lengths

Just as in the case of the up-down counter, the average run lengths are essentially determined by the threshold γ and the information number I, i.e., the average increment of the counter when the data are truly out-of-sync. For large γ the run lengths are approximately

$$\text{short ARL} \approx \frac{\gamma}{I}$$

and

$$\text{long ARL} \approx C\, e^{D\gamma},$$

where C, D, and I depend, of course, on k, the number of parity checks used in determining the counter increment. These approximations were borne out by the simulations reported in Section III, and the constants C and D were determined empirically for each k considered. I, the information number, is the average increment of the counter when the data are truly out-of-sync. (Of course, I depends upon k, the number of parity checks used in determining the counter increment, and is an increasing function of it.)

The information I from a pair of parity checks is the sum of the information number I_1 from the odd parity checks (the average increment of the counter at odd n) and I_2 from the even parity checks. Table 3 shows the contributions of these two parity check subsequences, revealing that the odd checks contribute roughly two-thirds of the total information.

In designing the detection algorithm, we must choose a symbol error probability p_e, at which the counter is designed to operate most efficiently. In practice, as the true symbol error probability varies with the signal-to-noise ratio, it will generally be smaller than p_e, so that fewer parity check failures occur. In this case, the counter will perform better whether the data is in sync or not; i.e., the short ARL will be reduced and the long ARL increased.

In case the signal-to-noise ratio is degraded so much that the symbol error probability exceeds p_e, however, a problem arises in the performance of the counter, as both the short and long ARL's become less favorable. In real life, the symbol error probability is not constant, and the fact that the long ARL's shorten during periods of low signal-to-noise ratio would introduce the same spurious loss of sync which seems to plague the current Viterbi decoders. If the short ARL gets longer during a period of high symbol error rate, this causes the response time to a loss of sync to increase and may cause a string of data to be lost *if* the short run hits at the time of low signal-to-noise ratio, but a shortening of the long ARL *whenever* there is such a period will have a far greater effect on the overall behavior of the system.

This degradation of long ARL can be totally eliminated by giving up the information I_2 from the even parity check subsequence. When the counter uses only the odd subsequence, the long ARL is unaffected by the signal-to-noise ratio. This is because the odd checks are completely random (independent, with 50% probability of success) whenever the process is in sync.

Since most of the information comes from the odd subsequence anyway, we believe it is prudent to base the counter on this subsequence alone. The performance parameters in the next section were all calculated for counters based solely on the out-of-sync parity checks.

III. Performance Numbers

The size of the ROM needed for counter increments is 2^k. As before, we will describe the counter in terms of the number k of parity checks needed to compute the counter increment.

Once k is chosen, the log-likelihood scheme determines the counter increments. The only question left in algorithm design is the threshold at which the system is declared out-of-sync. If the threshold is high, the probability of false loss of sync is low, or, equivalently, the time between false losses of sync is long. On the other hand, a high threshold will also make the short run length (or the time between loss of sync and detection of that loss of sync) large.

We first consider the influence of short run length on system performance. There is an obvious reason to want the short run length to be small: the sooner after a loss of sync that the system gets back on track, the better. But there is another reason as well. Voyager has a concatenated coding scheme: after the convolutional code is Viterbi decoded, an additional code, the Reed-Solomon code, is decoded. As far as the Reed-Solomon decoder is concerned, the data stream during the short run is just a stream of bad data. (Of course, if the out-of-sync condition was caused by the deletion of a symbol, and the node-synchronizer solves the problem by deleting another symbol, then a whole bit has been deleted, and frame boundaries are lost as well. In this case, the data during the short run can never be recovered. So we will consider the case in which the total number of channel symbols has not been changed. This will be true when the out-of-sync condition was caused by a spurious loss of sync caused by the node synchronizer, since the synchronizer will alternate adding and deleting channel symbols, and will be true half of the time anyway.) The decoder can recover a fair amount of bad data, and so if the short run is short enough, the Reed-Solomon decoder will be able to recover it most of the time. So the length of the average short run is not so important as the probability that the Reed-Solomon decoder will be able to recover the data in the short run. For the $k = 8$ counter, assuming Viterbi burst statistics for 2.3 dB (Ref. 8) and depth 4 interleaved Reed-Solomon words, Table 4 shows the probability of decoding for a word in a frame wholly containing a short run.

This same table shows the long run lengths (both in bits and in time, at the reasonable Voyager data rate of 20,000 bps) for these same thresholds. Looking, for example, at threshold 15, we see that the probability that a word contained in a short run will be corrected by the Reed-Solomon decoder is 2/3, and that sync is lost incorrectly every two days.

Can we do better? In fact, by going to $k = 16$, we can do much better. Table 5 shows this same information with $k = 16$. With $k = 16$ and threshold 14.5, a word which is in a frame attacked by a short run will decode correctly with probability 0.86, and the mean time between false losses of sync (average long run time) is 6.7 *years*.

These numbers were obtained by simulation methods explained in Appendix B.

Several other questions can be asked about the performance of the counter. What if a short run hits more than one frame? With $k = 16$ and threshold 14.5, the probability that a short run hits more than one frame is 0.033. And even if the short run does intersect more than one frame, the probability that it causes a word error in each frame is less than 0.005. Thus, the probability of a decoder error in each of two consecutive frames because of a spurious loss of node synchronization is less than 0.0002.

Another question is the probability of *some* loss of data due to a short run. We saw that in the case $k = 16$, threshold 14.5, the probability that any one word fails to decode is 0.14, but since decoding failures in the four words are by no means independent, this does not tell us the probability that there is some loss of data—that is, that one or more of the four interleaved Reed-Solomon words fails to decode. In the case $k = 16$, threshold 14.5, this probability is 0.19.

References

1. Deutsch, L. J., and Miller, R. L., "The Effects of Viterbi Decoder Node Synchronization Losses on the Telemetry Receiving System," *TDA Progress Report 42-68*, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 15, 1981.

2. Deutsch, L. J., and Miller, R. L., "Viterbi Decoder Node Synchronization Losses in the Reed-Solomon/Viterbi Concatenated Channel," *TDA Progress Report 42-71*, Jet Propulsion Laboratory, Pasadena, Calif., Nov. 15, 1982.

3. Liu, K. Y., and Lee, J. J., "An Experimental Study of the Concatenated Reed-Solomon/Viterbi Channel Coding System Performance and Its Impact on Space Communications," Publication 81-58, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 15, 1981.

4. Golomb, et al., *Digital Communications with Space Applications*, Prentice Hall, 1964, Chapter 7.

5. Greenhall, C. A., and Miller, R. L., "Design of a Quick-Look Decoder for the DSN (7,1/2) Convolutional Code," *DSN Progress Report 42-53*, Jet Propulsion Laboratory, Pasadena, Calif., Oct. 15, 1979.

6. Lorden, G., "Procedures for Reacting to a Change in Distribution," *Annals of Mathematical Statistics*, 42, No. 6, 1897-1908, 1971.

7. Feller, W., *An Introduction to Probability Theory and Its Applications*, John Wiley and Sons, 1968.

8. Miller, R. L., Deutsch, L. J., and Butman, S. A., On the Error Statistics of Viterbi Decoding and the Performance of Concatenated Codes, Publication 81-9, Jet Propulsion Laboratory, Pasadena, Calif., Sept. 1, 1981.

9. Siegmund, D., "Importance Sampling in the Monte Carlo Study of Sequential Tests," *Annals of Statistics*, 4, 673-684, 1976.

Table 1. Dependence of information numbers on symbol error probability

p	π	I_1	I_2	I
0.08	0.5875	0.0154	0.0155	0.0309
0.10	0.5537	5.78×10^{-3}	5.80×10^{-3}	0.0116
0.12	0.5321	2.07×10^{-3}	2.07×10^{-3}	4.14×10^{-3}

Table 2. Average run lengths for various thresholds

p	π	I	m	Short ARL	Long ARL
0.1	0.5537	0.01157	10	118	547
			15	205	2,062
			20	296	6,693
0.08	0.5875	0.03091	10	88	1,164
			15	145	7,496
			20	202	44,850

Table 3. Information numbers

k	Total	Out-of-sync	In-sync
16	0.0862	0.0601	0.0261
8	0.0633	0.0406	0.0228

Table 4. Counter performance with $k=8$

Threshold	Mean short run, bits	Probability that a Reed-Solomon word in a short run will decode	Mean long run	
			Bits	Time
5	116.6	0.98	6.1×10^4	3 seconds
10	238.6	0.86	5.2×10^7	43 minutes
15	363.3	0.66	4.5×10^{10}	26 days
16	387.6	0.62	1.7×10^{11}	100 days
18	438.0	0.52	2.6×10^{12}	4 years
20	487.4	0.43	3.8×10^{13}	60 years

Table 5. Counter performance with $k=16$

Threshold	Mean short run, bits	Probability that a Reed-Solomon word in a short run will decode	Mean long run	
			Bits	Time
5	88.6	0.99	5.9×10^6	5 minutes
10	171.4	0.95	7.9×10^9	4 days
11.5	196.4	0.93	6.0×10^{10}	35 days
13	221.7	0.90	5.0×10^{11}	290 days
14.5	246.9	0.87	4.2×10^{12}	6.7 years
18	305.1	0.77	6.1×10^{14}	960 years
20	338.8	0.71	1.0×10^{16}	16,000 years

Appendix A
Proofs

This appendix gives mathematical proofs of the statements made in Section II.

We use the same random processes to model the situation: (1) $\ldots, M_{-2}, M_0, M_2, \ldots$ an i.i.d. process, $P(M_0 = 0) = P(M_0 = 1) = 1/2$, representing the message; (2) $\ldots, E_{-2}, E_{-1}, E_0, E_1, \ldots$ an i.i.d. process, $P(E_0 = 1)$ = symbol error probability, representing the errors in the received sequence and independent of the sequence $\ldots, M_{-2}, M_0, M_1, \ldots$; (3) $\ldots, R_{-2}, R_{-1}, R_0, R_1, \ldots$, the convolutionally encoded M_i's added to the E_i's, representing the hard-quantized received channel symbols; and (4) $\ldots, X_{-1}, X_0, X_1, \ldots$ the sequence of parity checks derived from the R_i's.

Proposition: $p(X_n = x_n, X_{n-1} = x_{n-1}, \ldots, X_{n-2k} = x_{n-2k}) = [p(X_n = x_n, X_{n-2} = x_{n-2}, \ldots, X_{n-2k} = x_{n-2k}) \cdot p(X_{n-1} = x_{n-1}, \ldots, X_{n-2k+1} = x_{n-2k+1})]$.

Proof: First observe that for m odd,

$$p(X_m = 1 | X_{m-2}, X_{m-4}, \ldots, \ldots E_{-1}, E_0, E_1, \ldots) = 1/2,$$

because X_m is a sum of M_{m+1} and other variables, $p(M_{m+1} = 1) = 1/2$, and M_{m+1} is independent of all the random variables on which we are conditioning. This means that for odd m

$$p(X_m = x_m, X_{m-2} = x_{m-2}, \ldots,$$

$$X_{m-2\ell} = x_{m-2\ell} | \ldots, E_{-1}, E_0, E_1, \ldots) = 2^{-\ell-1} \quad \text{a.s.}$$

for any sequence $(x_m, x_{m-2}, \ldots, x_{m-2\ell})$ of zeroes and ones. But the values of the even parity checks are determined entirely by $\ldots, E_{-1}, E_0, E_1, \ldots$, and so are independent of the odd parity checks.

Notice: (1) The fact that we looked at a sequence of odd length was convenient for notation but had no effect on the proof. (2) We showed not only that the even and odd parity checks are independent, but that the odd parity checks are themselves i.i.d. with probability 1/2 of success. (3) The same result holds for the measure q, reversing the roles of even and odd.

Corollary: $p(X_n = x_n | X_{n-1}, X_{n-2}, \ldots) = p(X_n = x_n | X_{n-2}, X_{n-4}, \ldots)$, and the same for q. If n is odd, $p(X_n = x_n | X_{n-2}, X_{n-4}, \ldots) = 1/2$; for even n $q(X_n = x_n | X_{n-2}, \ldots) = 1/2$.

Definition: The one-sided, pseudo parity check sequence $\bar{X}_1, \bar{X}_2, \ldots$ is the parity check sequence based on the process $\ldots, 0, 0, R_1, R_2, R_3, \ldots$ That is, $\bar{X}_1 = R_1$, $\bar{X}_2 = R_2 + R_1$, $\bar{X}_3 = R_3 + R_2, \ldots, \bar{X}_n = X_n$ for $n \geqslant 14$.

Proposition: Every event of the form $(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$ corresponds to exactly one event $(R_1 = r_1, \ldots, R_n = r_n)$.

Proof: The X_i's are derived from the R_i's. Going backwards, knowing X_1 tells you R_1, and knowing X_i and $R_1, \ldots, R_{i-1}$ tells you R_i.

Corollary: A log-likelihood counter based on the X_i's will always contain exactly the same value as one based on the R_i's.

Proposition: If I_m is the information number of any log-likelihood counter with inputs based on the last m outputs of any discrete random process and I the information number with increments based on the indefinite past, then $\lim_{m \to \infty} I_m = I$.

Proof: For simplicity, we give the proof for random variables taking on the values zero and one. The proof for random variables taking on finitely many values is exactly the same.

Let $\ldots, Y_{-n}, \ldots, Y_0, Y_1, \ldots$ be the stationary process. (In our case, the Y_i's are pairs of hard-quantized received channel symbols). Let

$$f_n(y,j) = \log \left(\frac{[p(Y_0 = j \mid Y_{-1}, Y_{-2}, \ldots, Y_{-n})](y)}{[q(Y_0 = j \mid Y_{-1}, Y_{-2}, \ldots, Y_{-n})](y)} \right),$$

where y is chosen from the probability space on which the Y_i's are defined and j is zero or one.

It is a standard result of Martingale theory that

$$\lim_{n \to \infty} f_n(y,j) = f(y,j)$$

$$= \log \left(\frac{p(Y_0 = j | Y_{-1}, \cdots)(y)}{q(Y_0 = j |_{-1}, \cdots)(y)} \right), \text{ a.s.,}$$

and so by the dominated convergence theorem

$$\lim_{n\to\infty}\int f_n\,[y, Y_0(y)]\,dp = \int f\,[y, Y_0(y)]\,dp,$$

and the same for integration with respect to q. But these are just the information numbers for the counters.

Notice: A log-likelihood counter using k parity checks based on parity checks $\bar{X}$ will, after time $2k + 14$, have exactly the same increments as a log-likelihood counter using k parity checks and based on parity checks X, because the values of $\bar{X}$ and X are exactly the same starting at time 14.

Theorem: As k approaches infinity, the information in the log-likelihood counter whose increment at time n is

$$\log \frac{p\,(X_n = x_n \,|X_{n-2}, \ldots, X_{n-2k+2})}{q\,(X_n = x_n \,|X_{n-2}, \ldots, X_{n-2k+2})},$$

kept separately for n even and odd, approaches all the sync information in the hard-quantized received channel symbol stream $R_1, R_2, \ldots$

Proof: As m goes to infinity, the information in a counter $C(1)$ based on past of length $2k$ of the hard-quantized received channel stream approaches all the information in the stream. The values in $C(1)$ are exactly the same as would be in a counter $C(2)$ based on the past of length $2k$ of the (nonstationary) parity checks $\bar{X}_i$. For $n > 2k + 14$, the increments in counter $C(2)$ are the same as those in a counter $C(3)$ based on the probabilities of $(X_n \,|X_{n-1}, \ldots, X_{n-2k})$. But, since even and odd parity checks are independent, this is exactly the same as the counter of the theorem.

Appendix B
Simulation of ARL's

The performance figures of Section III were obtained by simulation. The short ARL's were simulated directly by generating independent symbol errors with probability p, computing the resulting parity check stream, and feeding it to the counter with threshold γ. Direct simulation of the long ARL's is not feasible, however, because (as the results show) the time required to generate a statistically useful sample of long run lengths would be too great.

The long ARL's were simulated by a modification of a standard technique called "importance sampling" (Ref. 9), in which the process to be analyzed – in this case, the parity check sequence – is generated using a probability distribution Q different from the distribution P for which results are desired. In our case, P specifies independent 50-50 results for the out-of-sync parity checks – so that the time for the counter to reach a distant threshold γ is quite large. A distribution Q was chosen to make the counter reach the threshold more quickly – namely, independent parity checks with probability of failure π^*, substantially less than 1/2.

The method of importance sampling is based on the simple fact that for any event, A, the probability of A under P can be obtained from simulations carried out under the distribution Q. The key is provided by the identity

$$P(A) = \int_A dP = \int_A \left(\frac{dP}{dQ}\right) dQ . \qquad \text{(B-1)}$$

Here the quantity dP/dQ is the Radon-Nikodym derivative of P with respect to Q, which in our application is simply the likelihood ratio

$$\frac{P(X_1)\cdots P(X_n)}{Q(X_1)\cdots Q(X_n)}$$

of the parity checks $X_1, \cdots, X_n$ up to the time that A occurs. Since $P(X_i) \equiv 1/2$, and $Q(X_i) = \pi^*$ if $X_i = 1$ (failure), $= 1 - \pi^*$ if $X_i = 0$ (success), relation (1) can be made more explicit. Using E_Q to denote expectation under Q, it takes the form

$$P(A) = E_Q\left[\left(\frac{1}{2\pi^*}\right)^F \left(\frac{1}{2(1-\pi^*)}\right)^S 1\{A\}\right], \qquad \text{(B-2)}$$

where F and S are the numbers of failures and successes, respectively, in the out-of-sync parity checks up to the time that A occurs, and $1\{A\} = 1$ if A occurs, $= 0$, otherwise.

The simulation of the long ARL was based on the definition of a *counter cycle*: Starting from a given state s^* of k zeroes and ones, the cycle ends the first time that the counter resets to zero with the same sequence s^* in its memory. Let T denote the time (number of symbol pairs) for a cycle to end and let N denote the number of cycles until the threshold γ is crossed. Then using a standard result called Wald's equation (Ref. 7, Vol. II, p. 567), we have

$$\text{long ARL} = EN \cdot ET .$$

(Actually, the right side gives the expected time until the *end* of the first cycle on which γ is crossed, but the extra time to end the cycle after crossing is negligible compared to the long ARL.) The quantity ET was easily simulated directly, since when the parity checks are random the cycles end fairly quickly (and don't depend on γ at all). The evaluation of EN was based on

$$EN = \frac{1}{P(A)},$$

where $A = \{\gamma \text{ is crossed}\}$ for a given cycle, and $P(A)$ was simulated using the method of importance sampling, as described above.

Importance sampling was used to estimate $P(A)$ in several independent sets of simulations. Not surprisingly, the estimates were more stable for smaller thresholds. After all the data were gathered, a least squares line was drawn through the points representing small thresholds (see Figs. B-1 and B-2). These lines were used for the long ARL's in Tables 4 and 5.

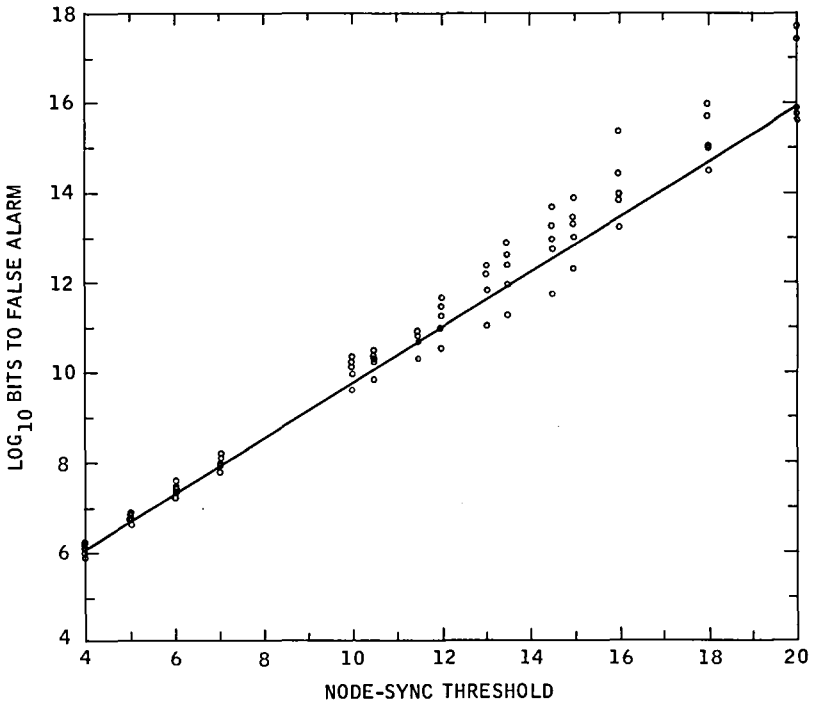


Fig. B-1. Mean time to false alarm, k=16; least squares line from thresholds 4, 5, 6, and 7

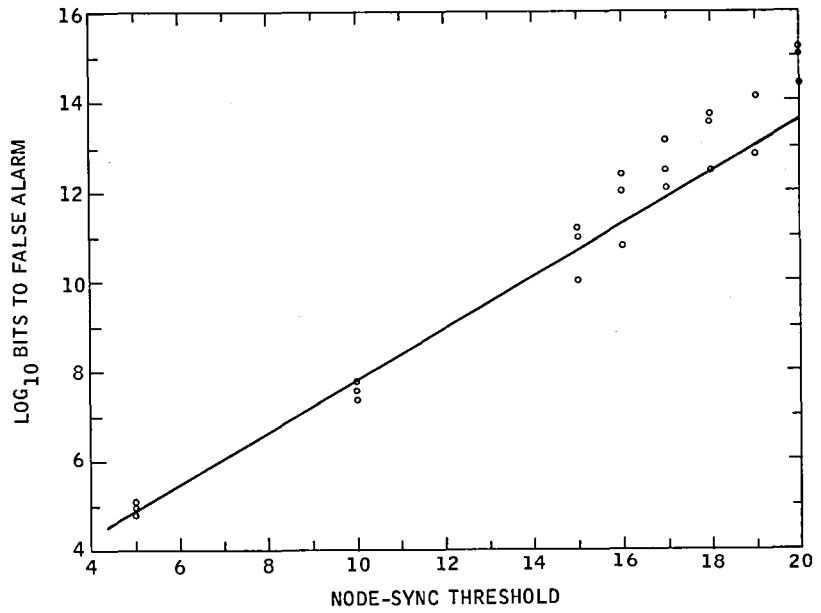


Fig. B-2. Mean time to false alarm, k=8; least squares line from thresholds 5, 10, and 15

Approximation to the Probability Density at the Output of a Photomultiplier Tube

R. J. Stokey and P. J. Lee
Communications Systems Research Section

In this paper, the probability density of the integrated output of a PMT is approximated by the Gaussian, Rayleigh, and Gamma probability densities. The accuracy of the approximations depends on the signal energy α: the Gamma distribution is accurate for all α, the Rayleigh distribution is accurate for small α ($\lesssim 1$ photon) and the Gaussian distribution is accurate for large α ($\gtrsim 10$ photons).

I. Introduction

H. H. Tan (Ref. 1) used a Markov diffusion model to determine an approximate probability density at the output of a photomultiplier tube (PMT). The resulting expression (Eq. 66, Ref. 1) is difficult to evaluate. Even by computer, the computational error can become serious since the number of significant terms in the summation of Bessel functions in Eq. (66) increases as α increases. Furthermore, in order to obtain the probabilities of detection and false alarm for a direct detection optical communications system (Ref. 2), this density must be convolved with a Gaussian noise density and integrated. As a result of these complications, three simply computed probability densities (Gaussian, Rayleigh, Gamma) will be considered as approximations to the exact density function.[1]

II. Statistical Model

The input to the PMT (Fig. 1) is a pulse with power λ photons/sec of duration T_s seconds. The number of photoelectrons emitted at the PMT cathode[2] during the time slot $[0, T_s]$ is assumed to be a poisson random variable with a mean value of $\alpha = \lambda T_s$.

The PMT is characterized by its average gain A and the number of dynode states ν. The variance of the PMT output is $2AB$ where $B = 1/2\ (A - 1)/(A^{1/\nu}-1)$. It is assumed that the bandwidth of the PMT (W_p) is greater than the integration bandwidth ($W_s = 1/T_s$); i.e., $W_p \gg W_s$, and consequently all the signal energy appears at the integrated output. The sampled voltage at the PMT output is $(eR/T_s)y$ where R is the anode resistance and $e = 1.6 \times 10^{-19}$ coulombs. The value y ($Y = y$) is a dimensionless "gain" and Y is the statistic of interest. The probability density of Y (Eq. 66, Ref. 1) is the combination of an impulse at $Y = 0$ and a non-impulse density, $P_X(y)$:

$$P_Y(y) = c\delta(y) + (1 - c)P_X(y) \quad ; \quad y \geq 0 \tag{1}$$

where

$$c = e^{\alpha(e^{-A/B}-1)} \tag{2}$$

[1] In this article, the density of Eq. (66), Ref. 1, will be called the exact density even though it is a Markov diffusion approximation of a Galton-Watson branching process.

[2] For the sake of simplicity in this paper, we assume that the PMT quantum efficiency is unity.

III. Approximations to the Non-Impulse Density $P_X(y)$

The mean and variance of the densities $P_Y(y)$ and $P_X(y)$ are

$$m_y = \alpha A \quad ; \quad m_x = m_y/(1-c) \tag{3}$$

$$\sigma_y^2 = \alpha(A^2 + 2AB) \quad ; \quad \sigma_x^2 = \frac{\sigma_y^2}{1-c} - cm_x^2 \quad . \tag{4}$$

A Rayleigh approximation to the probability density $P_X(y)$ is obtained by equating the mean values. The Gamma and Gaussian approximations are obtained by equating the first and second moments.

The three approximations to $P_X(y)$ and consequently $P_Y(y)$ (Eqs. 1, 2) are:

1. *Rayleigh Approximation*

$$P_X(y) \simeq \frac{y}{\sigma^2} \exp\left[-y^2/2\sigma^2\right] \quad ; \quad y \geqslant 0 \tag{5}$$

$$\sigma = \sqrt{\frac{2}{\pi}} m_x$$

2. *Gamma Approximation*

$$P_X(y) \simeq \frac{a}{\Gamma(b)} (ay)^{b-1} e^{-ay} \quad ; \quad y \geqslant 0$$

$$a = \frac{1}{k_1\left[\dfrac{k_2}{k_1^2} - 1\right]} \quad ; \quad b = ak_1 \tag{6}$$

$$k_1 = m_x \quad ; \quad k_2 = \frac{m_y^2 + \sigma_y^2}{1-c}$$

3. *Truncated Gaussian Approximation*

$$P_X(y) \simeq \frac{e^{\dfrac{-(y-m_x)^2}{2\sigma_x^2}}}{\sqrt{2\pi}\,\sigma_x\left[1 - Q\left(\dfrac{m_x}{\sigma_x}\right)\right]}; \quad y \geqslant 0 \tag{7}$$

$$Q(\alpha) = \frac{1}{\sqrt{2\pi}} \int_\alpha^\infty e^{-\beta^2/2}\, d\beta$$

IV. Discussion

The 11-dynode ($\nu = 11$) RCA 31034 A PMT has a usable average gain in a range from 10^5 to 10^7. The curves of Figs. 2-6 are calculated with an average gain of $A = 10^6$. Several different values of α are considered. The case of $\alpha = 5.2 \times 10^{-6}$ photons/slot (Fig. 2) corresponds to a typical dark current (T_s = 100 ns, λ = 52 counts/sec) in the optical channel (Ref. 2). For this case of small α (Fig. 2), the Rayleigh and Gamma distributions reproduce the shape of the exact distribution with the Gamma distribution being more accurate in the tail regions. Figures 3-5 demonstrate some representative light current values for the optical channel. The fidelity of the Gamma approximation varies as a function of α and is most accurate for $\alpha \simeq 1$ and becomes asymptotically close as α increases. Figure 6 demonstrates the validity of the Gaussian approximation for large α. For the Rayleigh distribution, the ratio of the mean and standard deviation is equal to a constant since the density function depends on one parameter. Consequently, the Rayleigh distribution becomes a poor approximation as the mean value increases. Figures 2-6 demonstrate the accuracy of the Gamma distribution for all values of α. The Rayleigh and Gaussian distributions are easily computed functions which can be used for small and large values of α respectively.

References

1. Tan, H. H., "A Statistical Model of the Photomultiplier Gain Process with Applications to Optical Pulse Detection," *TDA Progress Report 42-68*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 55-67, Feb. 15, 1982.

2. Lesh, J., Katz, J., Tan, H., and Zwillinger, D., "The 2.5 Bit/Detected Photon Demonstration Program: Description, Analysis and Phase I. Results," *TDA Progress Report 42-66*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 115-132, Dec. 15, 1981.

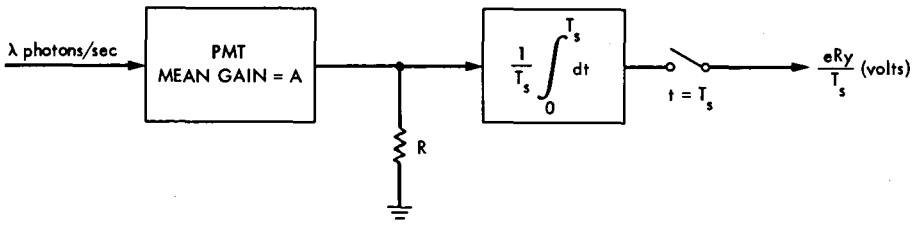


Fig. 1. Model of the integrated output of a random gain PMT

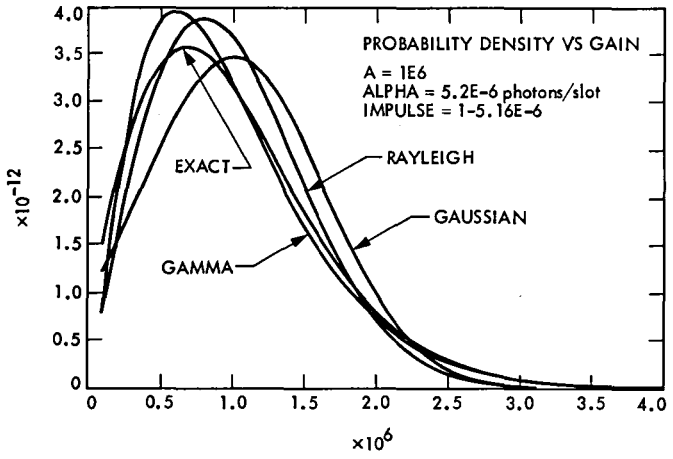


Fig. 2. Small α case: A typical dark current of λ = 52 cts/s will result in $\alpha = 5.2 \times 10^{-6}$ photons/slot when T_s = 100 ns. For this case, the probability that the output is nonzero is 5.16×10^{-6}

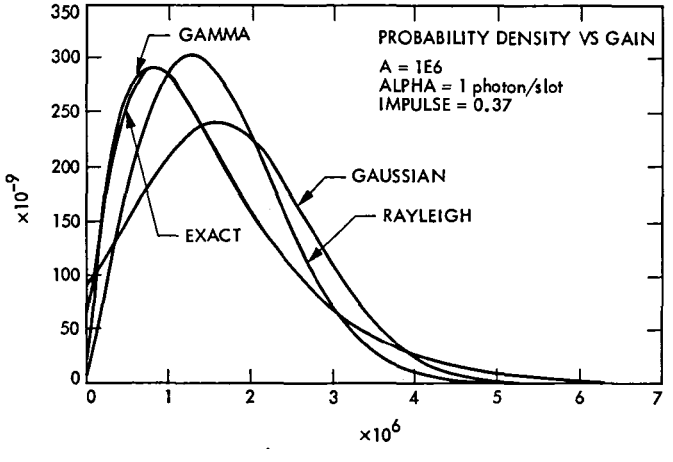


Fig. 3. Moderate α case: An optical channel (256-ary PPM) light current of α = 1 photon/slot corresponds to 8 bits/photon; i.e., $\rho = 8/\alpha$ (bits/photon)

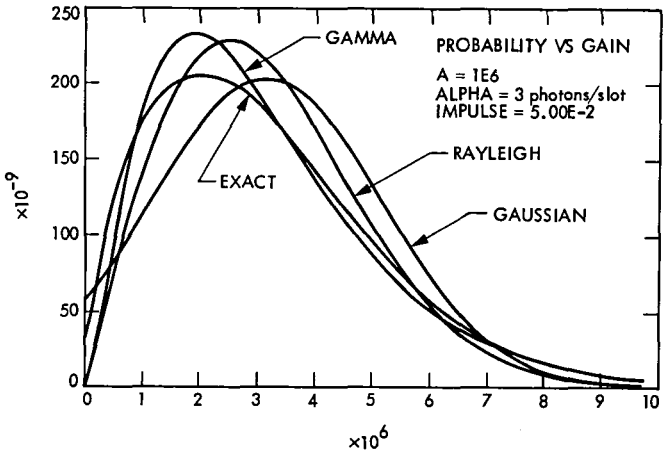


Fig. 4. Moderate α case: An optical channel (256-ary PPM) light current of $\alpha = 3$ photons/slot corresponds to 2.67 bits/photons; i.e., $\rho = 8/\alpha$ (bits/photon)

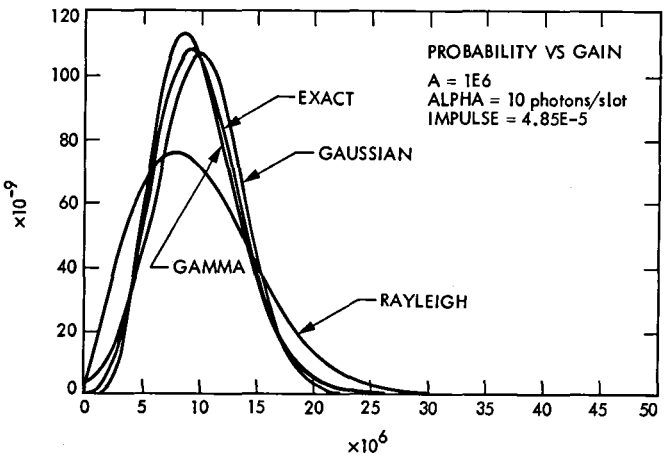


Fig. 5. Moderate α case: An optical channel (256-ary PPM) light current of $\alpha = 10$ photons/slot corresponds to .8 bits/photons; i.e., $\rho = 8/\alpha$ (bits/photon)

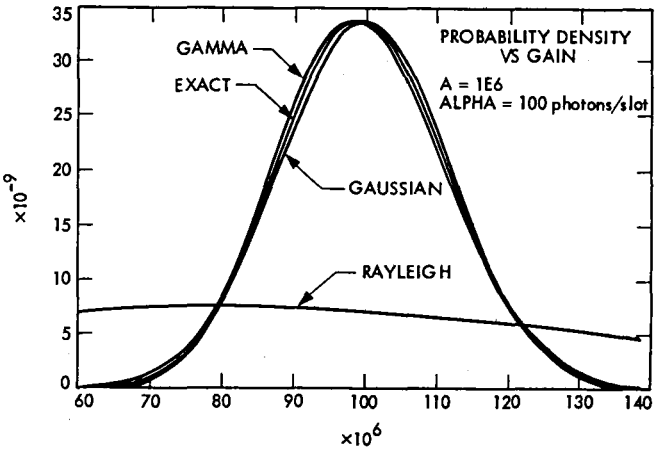


Fig. 6. Large α case: An optical channel (256-ary PPM) light channel of $\alpha = 100$ photons/slot corresponds to $\rho = 12.5$ photons/bit

Unified Analysis for Antenna Pointing and Structural Errors Part I. Review

K. Abichandani
Ground Antennas and Facilities Engineering Section

A necessary step in the design of a high-accuracy microwave antenna system is to establish the signal error budget due to structural, pointing, and environmental parameters. This report, the first of a series, deals with a unified approach in performing error budget analysis as applicable to ground-based microwave antennas of different size and operating frequency.

A discussion of major error sources contributing to the resultant deviation in antenna boresighting in pointing and tracking modes and the derivation of the governing equations are presented. Two computer programs (SAMCON and EBAP) have been developed in-house, including the antenna servo-control program, as valuable tools in the error budget determination. A list of possible errors giving their relative contributions and levels is presented.

I. Introduction

The primary mission of a microwave antenna system is to receive signals from a deep space target either by automatic tracking or by pointing at the target in response to a predetermined command pattern. Both the accuracy in pointing the microwave beam and the precise automatic tracking become imperative and require a careful investigation of the bounds of each (Refs. 1 to 14). As the antenna size and the communication frequency increase, the size of errors and the resulting gain loss take on additional importance.

The practice by microwave and structure engineers in designing new microwave antennas is readily available in the literature. Antennas of different size, operating frequency, and mount have been designed and constructed in the past in different countries. Their design specifications of the allowable deviations from a true theoretical behavior have been assumed based on past experience. Also, there has been a lack of sufficient theoretical background for error predictions based on a unified approach to quantify the errors prior to antenna construction.

This report attempts to bridge the gap by outlining, in a unified approach, methods to determine the contribution by various structural/pointing sources of error on the overall signal gain loss during antenna pointing and/or tracking

regardless of the antenna configuration, size, operating frequency, etc.

II. Analysis

Errors may be classified broadly into two principal categories: dependent errors and independent errors. Dependent errors may be grouped together to form independent sets. Examples of dependent error sources are errors due to wind on the main reflector, supports, feed cone intermediate reference assembly, and the instrument tower. Each independent set, and individually independent error sources, may be applied to appropriate input points of the system for evaluation. To clarify this point, Fig. 1 depicts the antenna control system block diagram for one axis movement of the antenna and indicates likely error sources. Their contribution to the total signal error naturally depends upon positioning. To illustrate this effect, let us consider the simplified representation depicted in Fig. 2 where G_1, G_2, and H are the Laplace transforms for a general closed-loop control system components that could represent an antenna control system. For example, R is the controlled quantity (e.g., antenna position in elevation angle, declination angle, or azimuth angle), C is the command signal, and E is the error between the command and generally modified (or shaped) controlled quantity R. If S_i ($i = 1, 2, 3$) are the error sources applied at the components G_1, G_2, and H, one may write, in general,

$$\frac{R}{C} = \frac{G_1 G_2}{1 + HG_1 G_2} \tag{1}$$

$$\frac{E}{C} = \frac{1}{1 + HG_1 G_2} \tag{2}$$

$$\frac{R}{S_1} = \frac{G_1 G_2}{1 + HG_1 G_2} \tag{3}$$

$$\frac{R}{S_2} = \frac{G_2}{1 + HG_1 G_2} \tag{4}$$

$$\frac{R}{S_3} = \frac{HG_1 G_2}{1 + HG_1 G_2} \tag{5}$$

Equations (1) through (5) are general for the control system shown and will be used in determining errors in antenna pointing and tracking.

III. Error Sources and Points of Application

Ten major error sources were identified and grouped as follows:

(1) Command errors.

(2) Command quantizing errors.

(3) Structure dead load errors.

(4) Wind-induced errors.

  (a) Steady wind errors.

  (b) Random wind effect.

  (c) Turbulent wind component.

(5) Data system errors.

(6) Tracking errors.

(7) Thermal errors.

(8) Structural tolerances and alignment errors.

(9) Servo control system errors.

(10) Refraction, microseism and instrument errors.

Each error source is discussed in detail as follows:

A. Command Errors

Commands are generally given according to predetermined positioning and rate schedule, e.g., a position command, or a tracking command. If an error is made in the predetermination of the commands, an error is naturally introduced at the outset. The error ΔC will be attenuated as follows: if ΔR is the change in the controlled quantity, then from Eq. (1),

$$\Delta R = \Delta C \frac{G_1 G_2}{1 + HG_1 G_2} \tag{6}$$

B. Command Quantizing Errors

In the case of an antenna tracking a space vehicle, the vehicle orbit is known a priori. By computation on a digital computer, the orbit configuration data are stored and the command signals are derived from this stored information in binary form. The quantizing error usually has a distribution of one bit. This amounts to 1 arcsecond rms in azimuth angle for a 20-bit quantization of the encoder.

C. Structure Dead Load Errors

Forces due to gravity resulting in moments on the structure deflect the members and panels, causing a change of shape of the antenna. As the elevation varies, so do the moments of

forces giving rise to structural deflections. However, the moments of forces are predictable (Refs. 8 to 14) and repeatable within the hysteresis characteristics of the structure. These characteristics can be determined a priori and stored by a computer program and thus calibrated out rather accurately. The compensation due to this effect will be discussed later in Section IV.

D. Wind-Induced Errors

The wind effect may be decomposed into two portions: first the portion due to a steady wind, second the portion due to a variable wind comprising gusts and direction changes. The two components of wind produce forces and moments on the antenna structure which can be superimposed and result in the boresight errors. Each portion is discussed below.

1. Steady wind errors. Steady wind may be analyzed by considering the aerodynamics of its flow on structures. The laminar or low Reynolds number component causes drag and lift forces that cause steady moments, depending on the location of the center of pressure.

The steady wind also has turbulent component and vortices associated with it. Turbulent flow results in high frequency variations, which the inertia of the antenna structure filters out. Vortex flow results in low frequency variations that require further consideration. These variations are amenable to power density analysis. If the "pure" steady-state wind velocity is V, the associated moment M may be computed as

$$M = \frac{\rho}{2} V^2 S \bar{c} C_m \tag{7}$$

where

ρ = density of air at sea level

V = wind velocity

S = projected area of the antenna dish

C_m = moment coefficient

$\bar{c}$ = mean diameter of the antenna

Once C_m is established, the moment can be computed. Knowing the stiffness of the structure, the deflection can be determined. The effect of "purely" steady wind from Eq. (7) can consequently be compared to that due to the gravity loading discussed above.

The effects of wind-induced stresses on the antenna structure have been addressed in detail in Ref. 8.

2. Random wind effect. In order to evaluate the effect of the random wind on the antenna structure (input) we will use the method which determines the mean square error in the antenna rigid body dynamics (output) resulting from a finite torque power spectrum. Let the power spectral density (PSD) of the variable wind distribution be denoted by $S_w(jw)$, and the general linear system transfer function of the antenna by $Y(jw)$; the mean square error $\overline{\theta}^2$ of the controlled quantity may be expressed as (Ref. 15)

$$\overline{\theta}^2 = \int_{-\infty}^{\infty} |Y(jw)|^2 S_w(jw)\, dw \tag{8}$$

where $Y(jw)$ is a general term representing either the dynamics of the antenna control system or the flexible dynamics of the antenna structure. The antenna servo-control simulation program, SAMCON, initiated by Ref. 16 performs, among others, the operation represented by Eq. (8) for the computation of the random wind effects.

3. Turbulent wind component. The power spectral density of the wind acting on the supporting structure is not largely different from that acting on the antenna main reflector. The only difference will be in the magnitude of the PSD of the torque due to the wind. The transfer function $Y(jw)$ associated with the supporting structure will have a form similar to that for the antenna dynamics, i.e., a second-order system transfer function represented by the following expression:

$$Y(jw) = \left(\frac{a_2 S^2 + a_1 S + a_0}{S^2 + 2\zeta W_n S + W_n^2} \right)_{s = jw} \tag{9}$$

where a_0, a_1 and a_2 are constants, ζ is the structural damping ratio, and W_n is the structural natural frequency of the mode selected for the analysis. The selection criterion considers the bandwidth of the antenna servo-control loop in the case of the antenna. In the case of the supporting structure, its modal frequencies within or near the wind spectrum may be examined.

The following widely accepted expression derived in Ref. (17) may be used for the wind PSD used in Eq. (8).

$$S_w(jw) = \frac{K_w}{(w_0^2 + w^2)(w_1^2 + w^2)} \tag{10}$$

where

S_w = wind torque PSD

K_w = constant establishing torque level which is written as

$$K_w = 4\alpha^2 w_0 w_1 (w_0 + w_1) T_0^2/\pi \qquad (11)$$

where

α = rms turbulence wind speed variation about the mean wind speed T_0: typically = 0.2

w = frequency in radians/sec

w_0 = frequency at first "break corner," radians/sec

E. Data Encoding/Displaying System Errors

If the antenna position and rate are remotely monitored in a central control room, the fine instruments, measuring that antenna position or rate including pinions, encoders, display, etc., may indicate erroneous readings due to errors inherent in the measuring instrument itself. One way to determine such errors is by statistically determined test data over a certain period of time. Sample values of typical error sources in antenna components are presented in Table 1 for information.

F. Track Errors

Tracking error can be defined as the space angle difference between the communication RF axis of the antenna and the vector to the RF source being tracked. The servo-control simulation program, SAMCON, which is being developed inhouse, provides the tracking error magnitude. One may choose arbitrarily the pickoff point for the position feedback which affects antenna structural flexure influence on the tracking errors. If data system gears are used, then the flexure effects will be circumvented. Additional details about the inputs and outputs for this program will be addressed in detail in a separate publication.

G. Thermal Errors

Another major source of error is that due to the temperature gradient in the various members of the antenna structure. Although the effect of the temperature gradient on an individual member can be predicted and accurately determined, determination of the combined effect of all the members comprising the antenna structure is a complex task. One approach would be to develop a computerized thermal model to determine the deformations and the resultant antenna errors. However, in the absence of such a model at present, field data from existing antennas may be used (Ref. 18). The main source of thermal error may be assumed to be due to the supporting structure, e.g., quadripods and the reflector supports. A pattern of temperature gradient from this data may be assumed; however, this procedure is approximate. Sometimes the thermal effects may be accounted for by using an empirical correction factor > 1.0 to be applied to the errors computed due to the dead load. The development of a good thermal model is now in progress (Ref. 19) to satisfy this need.

H. Structural Tolerances and Alignment Errors

This alignment error category may be divided into five major areas comprising:

(1) Alignment of radio frequency boresight axis to an Intermediate Assembly (IRA) (wherever existing).

(2) Alignment of an IRA to elevation axis, if IRA is provided.

(3) Alignment of elevation axis to azimuth axis.

(4) Alignment of azimuth axis to earth reference system.

(5) Runout of azimuth and elevation drive bearings.

The alignment of the RF boresight axis to an IRA is the most crucial one. The error consists of field alignment of the hyperboloid subreflector and its supports, field setting and alignment of the paraboloidal reflector and its supports including face panels, field alignment of the feed cone, and field alignment of the IRA. Contributions from the individual sources to the total boresight shift make up one large error which is part of the RF boresighting problem to be addressed later.

The RF boresighting problem involves the method of making the pattern measurement and boresighting the RF axis to the mechanical axis represented by the IRA. This problem consists of accurately locating a source of RF energy which can be moved throughout the hemisphere of coverage at ranges great enough to include the atmospheric refraction of the RF energy. Several methods have been considered for obtaining the location of the boresight axis relative to some reference point in space. The method with good promise is to arrive at the boresight error by CONSCAN techniques of pointing on a radio star (Ref. 20). However, this error can be determined rather accurately and corrected for by properly aligning the structure from the RF boresight to the IRA; the design of the data system eliminates the remaining structural manufacturing and alignment errors from the system output. The four errors that can thus be eliminated include (1) alignment of the IRA to elevation axis, (2) alignment of elevation to the azimuth axis, (3) alignment of the azimuth axis to an earth reference system, and (4) bearing runout.

The total effect of the above errors (derived in Ref. 21) is given in the Appendix and results in the following two equations.

$$\text{Error}(C)_x = -S_1 + C_1\,[\cos(h - C_2)]\left(\frac{\partial X}{\partial h}\right)_\psi - S_2 \tan Y \sin X + S_3 \tan Y \cos X - S_4 \tan Y + \frac{C_3}{\cos Y} \pm S_5 + A_3 \ddot{X} - S_7\left(\frac{\partial X}{\partial \lambda}\right)_\phi + A_2 \frac{X}{|X|} + A_5 \frac{\dot{X}}{|\dot{X}|} + S_8\left(\frac{\partial X}{\partial \phi}\right)_\lambda + A_7 \dot{X} + A_9 \frac{\sin 2\theta}{\cos Y} \qquad (12)$$

$$\text{Error}(C)_y = C_6 + C_4\,[\cos(h - C_5)]\left(\frac{\partial Y}{\partial h}\right)_\psi - S_2 \cos X - S_3 \sin X \pm S_6 + A_1 \ddot{Y} - S_7\left(\frac{\partial Y}{\partial \lambda}\right)_\phi + A_4 \frac{Y}{|Y|} + A_6 \frac{\dot{Y}}{|\dot{Y}|} + S_8\left(\frac{\partial Y}{\partial \phi}\right)_\lambda + A_8 \dot{Y} + A_{10} \sin 2\theta + A_{11} \cos 2\theta \qquad (13)$$

where (X, Y) is generalized position coordinates, $(\dot{X}, \dot{Y})$ is the velocity, $(\ddot{X}, \ddot{Y})$ is the acceleration, A and C are coefficients to be determined by field measurements, and S's are biases. Although the equations were originally derived for a 9-meter antenna operating in the S-band frequency (Ref. 21), their form and the technique used for evaluating the coefficients make them suitable without significant modifications for any size antenna system at any frequency. This is simply due to the fact that in the derivation followed there are no size- or frequency-dependent terms involved.

Equations (12) and (13) are algebraic summations of systematic RF to optical boresight error and optical boresight to true encoder axis error. The parameters X, Y, A, C, and S are listed in the definition of symbols table. Although the derivation of error equations was made with respect to two orthogonal axes, X and Y axes of the antenna, the analysis is still in general form and could be used for any two orthogonal coordinates. The only conditions to be satisfied are that (1) the axes be orthogonal, and (2) the chosen antenna axes relationship with these two orthogonal axes must be known a priori.

The five error sources that are included in the derivation of Eqs. (12) and (13) are

(1) X-Y antenna angular errors due to tilt of the primary (X) axis.

(2) X-Y antenna angular errors due to lack of orthogonality.

(3) Pointing error due to positioning errors in station location such as latitude and longitude of the antenna position.

(4) Pointing errors due to structural deflection.

(5) Boresight shift with polarization rotation.

The polarization rotation concerns unified S-band antennas and aircraft which were used to draw tracks providing linearly polarized transmission signals for measuring this error source (applicable to 9- to 26-meter antennas, Ref. 21). This error source is not applicable to nonpolarized signal transmission.

Equations (12) and (13) were programmed in-house in the Error Budget Analysis program (EBAP). In order to be able to determine the errors $(C)_x$ and $(C)_y$, all the coefficients in Eqs. (12) and (13) must be known a priori.

I. Servo Control System Errors

The error sources contributing to the positioning error of the antenna are:

(1) The type of the control system.

(2) Hardware.

(3) Tachometer.

The type of control system could be either type-1 or type-2. The error will be governed by the type of command to the particular type of system. For a type-2 system, a constant command will result in zero error; so will a constant rate command. Second, hardware such as motors, amplifiers, modulators, demodulators, synchros, etc., may have dead zones, thresholds, and tooth ripples which result in nonlinear operation near the theoretical zero error zones, such as final commanded position. Third, the function of the tachometer is to augment damping in a rate loop. Inaccuracy in the scale factor, presence of noise, and ripple will introduce error not only in following the rate commands, but also in the desired position because of effective nonlinear damping. Knowing the system dynamic characteristics from manufacturer's data, the servo simulation program (SAMCON), Ref. 16, provides the error magnitudes due to this source. Thus, all the errors in the control system may be determined.

J. Refraction, Microseism and Instrument Tower Errors

These sources of error have been reported to be insignificant. However, they are mentioned here briefly.

Refraction is an integral part of the boresight procedure which has been discussed earlier in Section III-H.

Microseisms, in general, consist of small amplitude, relatively high-frequency earth tremors. Microseisms can introduce errors into the system by disturbing the instrument tower. Reference 22 concluded that the largest error on the instrument tower caused by all possible sources was due to 96 km/hr (60 mph) wind blowing on the main reflector and transmitting overturning forces from the main pedestal through the earth to the instrument tower. This wind-induced error swamped out any errors due to microseisms. The magnitude of this error was about 0.3 arcseconds peak. Hence, microseisms are negligible compared to other errors in the system.

Having discussed the various error sources and the derivation of formulas to compute these, a list is given in Table 1 with some sample values.

IV. Relevant Work

In addition to the above pointing and tracking errors determination, other in-house work was initiated for determination of errors due to any reflector shape change, the consequent shift in the focal point of the main reflector, and the feed system components for a Cassegrain antenna. The pointing error sources were as follows (Ref. 23):

(1) Antenna primary reflector parabola translation.

(2) Primary reflector parabola rotation.

(3) Subreflector hyperbola translation.

(4) Hyperbola rotation.

Effect of wind load was included by applying a correction factor. This correction factor is the square of the ratio of two values of wind velocity: (1) the wind velocity for which the parameters are known a priori, and (2) the wind velocity for which the errors are to be computed.

The pointing error is computed as the root sum square (rss) of errors about two orthogonal axes in seconds of arc. This error source is root-sum-squared with the other sources discussed earlier. Changes in the shape of the antenna due to its dead weight and wind must be accounted for if errors are to be computed accurately.

V. Conclusions

Two computer programs (SAMCON and EBAP) were developed as analytical tools to support the error analysis study. Error equations for the EBAP program are listed in the Appendix. These programs will enable one to compute an error budget covering most of the sources outlined above.

Two major error sources emerged as requiring further development: one is the contribution by structural dynamics excited by some control commands and noise at various points of the antenna servo control system; the other is that due to thermal gradient. The first requires expanding the control simulation program (SAMCON) while the thermal gradient effect will require additional experimentation and analytical development.

In evaluating the errors given by Eqs. (12) and (13), some coefficients will have to be determined. By using known data from currently existing designs, preliminary error values may be computed. Updating the information after constructing any antenna will provide a better baseline for error determination of new ones. This is an iterative process which is important for designing new antennas in the future.

Acknowledgement

The author is indebted to R. Levy, M. Katow, H. Phillips, F. Lansing and I. Khan of the JPL Technical Staff for sharing their technical views and their editing contributions throughout this work.

References

1. Ruze, J., "Lateral Feed Displacement in a Paraboloid," *IEEE Trans. Ant. Prop.*, Sept. 1965, pp. 660-665, 1965.

2. Mar, J. W., and Liebowitz, J., Editors, *Structures Technology for Large Radio and Radar Telescope Systems*, The MIT Press, Massachusetts Institute of Technology, Cambridge, Mass., 1969, pp. 1-54.

3. Ibid, pp. 65-87.

4. Ibid, pp. 101-134.

5. Ibid, pp. 185-200.

6. Ibid, pp. 499-514.

7. Katow, M. S., Bartos, K. P., and Matsumoto, L. R., *JPL Modified STAIR Systems Operating Procedures*, NASA TM 33-304, Jet Propulsion Laboratory, Pasadena, Calif., June 1, 1968.

8. Levy, R., and Katow, M. S., "Implementation of Wind Performance Studies for Large Antenna Structures," *IEEE Mechanical Engineering in Radar*, pp. 27-33.

9. Levy, R., and Parzynski, W., "Optimality Criteria Solution Strategies in Multiple Constraint Design Optimization," 22nd Structures, Structural Dynamics and Materials Conference, Atlanta, Georgia, Apr. 6, 1981.

10. Levy, R., "Optimality Criteria Design and Stress Constraint Processing," Preprint, International Symposium on Optimum Structural Design, University of Arizona, Tucson, Oct. 19, 1981.

11. Katow, M. S., "Structural Design of a 64-Meter Low Cost Antenna," *DSN Progress Report 42-45*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 258-275, June 15, 1978.

12. Levy, R., *Iterative Design of Antenna Structures*, Technical Report 32-1526, Vol. XII, Jet Propulsion Laboratory, Pasadena, Calif., pp. 100-111, Dec. 15, 1972.

13. Levy, R., "Conceptual Studies for New Low Cost 64-m Antennas," *DSN Progress Report 42-33*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 55-67, June 15, 1976.

14. Katow, M. S., and Patton, R. H., "Structural Data Checks with Computer Graphics," *DSN Progress Report 42-27*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 148-153, June 15, 1975.

15. Davenport, W. B., Jr., and Root, W. L., *An Introduction to the Theory of Random Signals and Noise*, McGraw-Hill Book Company, Inc., N.Y., 1958.

16. James, M. R., "The Implementation and Analysis of a Baseline Antenna Servo Control System Model," Report by Harris Corporation, Melbourne, Florida, for Jet Propulsion Laboratory, NASA Contract No. NAS7-100, Subcontract No. 956125, Apr. 1982.

17. Sweo, E.A., "Method of Calculation for Reflector Assembly Pointing Error," Dalmo Victor Co., Report DV-230-1B, Belmont, Calif., Oct. 1963.

18. Glazer, S., and Gale, G., "Thermal Measurement Technique of Rib Elements on DSN Antenna Structure," *TDA Progress Report 42-66*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 67-79, Dec. 15, 1981.

19. Schonfeld, D., and Lansing, F., "Thermal Analysis of Antenna Backup Structure, Part I – Methodology Development," *TDA Progress Report 42-70*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 110-116, Aug. 15, 1982.

20. Abichandani, K. P., and Ohlson, J. E., "Boresighting Techniques for the Antenna Control Assembly (ACA)," *TDA Progress Report 42-64*, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 15, 1981.

21. Collins Radio Company, "Apollo Unified S-Band GSFC 30-Foot Antenna System, Kennedy Station," Engineering Report No. 1523=0559008-001 D3m, prepared for NASA Goddard Space Flight Center, Greenbelt, Md., Oct. 1966.

22. Isber, A. M., "Obtaining Beam-Pointing Accuracy with Cassegrain Antennas," *Microwaves*, pp. 40–44, Aug. 1967.

Definition of Symbols

A_1 Y acceleration error coefficient

A_2 X zenith structure shift

A_3 X acceleration error coefficient

A_4 Y zenith structure shift

A_5 X direction effect

A_6 Y direction effect

A_7 X velocity error coefficient

A_8 Y velocity error coefficient

A_9 X polarization shift component

A_{10} Y polarization shift component

A_{11} Y polarization shift component

C_1 X deflection coefficient in an elevation plane

C_2 X deflection angular offset in an elevation plane

C_3 Rf axis to Y axis lack of orthogonality (a positive sign means $|C_7|$ is subtracted from 90°) $\equiv \delta R$

C_4 Y deflection coefficient in an elevation plane

C_5 Y deflection angular offset in an elevation plane

C_6 RF axis to Y encoder axis bias

h elevation angle

S_1 X encoder bias

S_2 tilt component δV of the X axis, taken from a reference direction

S_3 tilt component δE of the X axis

S_4 X axis to Y axis lack of orthogonality (a positive sign means $|S|$ is subtracted from 90°) $\equiv \delta N$

S_5 X encoder hysteresis

S_6 Y encoder hysteresis

S_7 station longitude error ($\Delta\lambda$)

S_8 station latitude error ($\Delta\phi$)

t local hour angle

α, β, γ direction angles of target vector w.r.t. fixed frame

δ declination angle/infinitesimal change

θ the angle between a reference line on the antenna feed which is vertical when XY mount faces the east or south (depending on the antenna design) horizon and a constant vertical line.

λ station longitude angle

ϕ station latitude angle

ψ azimuth angle

Table 1. Sample values of some error sources

Type of error	Sample values (3σ) (for 64-m antenna), arcsec
Command error	
Input command determination	24
Input quantization	4
Command	7
Velocity	7
Acceleration	2
Friction	3
Tooth cogging and drive serve	15
Dead load errors	
Reflector	120
Hyperboloid translation	−14
Hyperboloid rotation	7
Intermediate reference assembly	21
Instrument mount	3
Wind-induced errors	
Steady wind (50 km/h)	5
Variable wind	25
Thermal errors	50
Structural tolerances and alignment errors	36
Data system errors	12
Alignment of normal to mirror surface parallel to normal to IRA	–
Optical path	
Alignment of 2-axis autocollimator to its support	–
Orthogonality between axes	–
Alignment of encoders	–
Encoder positional errors	–
Alignment of instrument mount to the gravity vector	–
Errors in digital computer and D/A converter	–
Errors in servo control system	–
Hardware	–
System type	–
Refraction, microseism	Less than 3

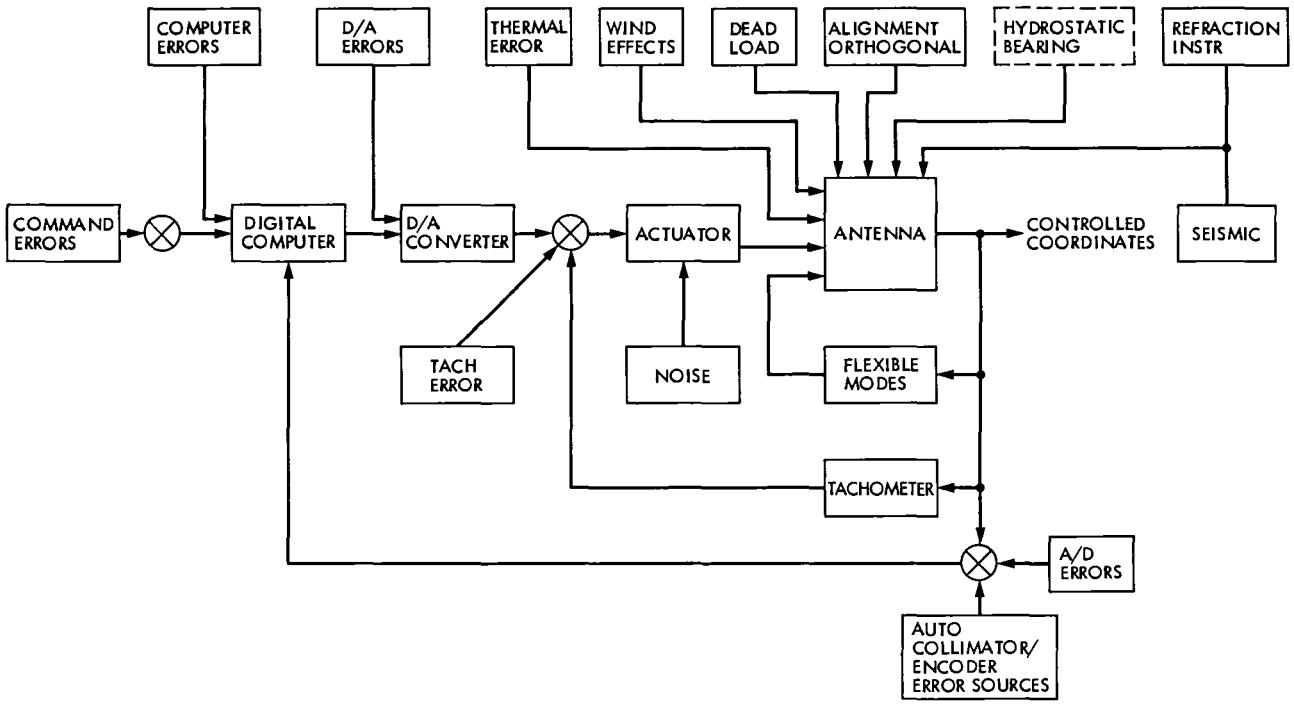


Fig. 1. Block diagram of error-source simulator

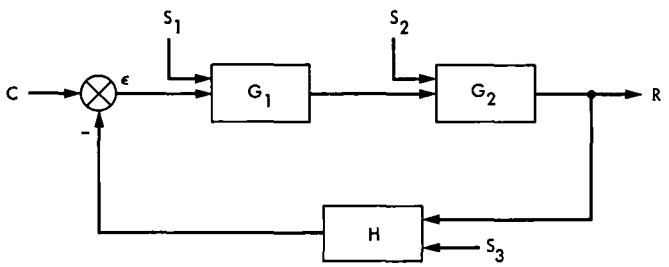


Fig. 2. Block diagram showing possible relative positions of error sources and control system components

Appendix

Error Equation Derivation

This appendix consists of derivations of several terms of the RF-to-encoder error equations. Spherical triangles are used since they display the error effects most vividly.

I. Antenna Angular Errors Due to Tilt of the Primary Axis

The two errors in X and Y angles, ΔX and ΔY, derived below[1], are those that occur as a result of successive rotation of the primary X axis of the antenna about each of the other two orthogonal axes. For small rotations, these are equivalent to two tilt components of the X axis. The following equations for ΔX and ΔY are used to solve for the axis misalignment or tilt components, δ_U and δ_E. The corrections, ΔX and ΔY, due to misalignment components δ_U, δ_E will be proved later as

$$\Delta X = \tan Y \, [\delta_U \sin X - \delta_E \cos X] \tag{A-1}$$

$$\Delta Y = \delta_E \sin X + \delta_U \cos X \tag{A-2}$$

where

$$\begin{aligned} X &= X' - \Delta X \\ Y &= Y' - \Delta Y \end{aligned} \tag{A-3}$$

In order to prove Eqs. (A-1) and (A-2) refer to Fig. A-1. The dotted tilted axes are obtained from the solid aligned axes (1), (2), and (3) by two steps:

(a) Rotate (2) by angle δ_U about axis (1) counterclockwise.

(b) Then rotate by angle δ_E about axis (3) clockwise.

Both components x, y, and z along axes (1), (2), and (3), and direction cosines $\cos\alpha$, $\cos\beta$ and $\cos\gamma$ of a target position vector have the same relations as follows, where primed values refer to tilted axes and unprimed values refer to aligned axes.

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = [BA] \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

and the new direction cosines are obtained from the unprimed direction cosines as

$$\begin{bmatrix} \cos\alpha' \\ \cos\beta' \\ \cos\gamma' \end{bmatrix} = [BA] \begin{bmatrix} \cos\alpha \\ \cos\beta \\ \cos\gamma \end{bmatrix} \tag{A-4}$$

where $[A]$ is the matrix obtained from step one and $[B]$ is the matrix obtained from step (b), above. $[BA]$ may be written as:

$$[BA] = \begin{bmatrix} \cos\delta_E & -\sin\delta_E & 0 \\ \sin\delta_E & \cos\delta_E & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\delta_u & \sin\delta_U \\ 0 & -\sin\delta_U & \cos\delta_U \end{bmatrix}$$

For small δ_E, δ_U, then, $\cos\delta_E \to 1$, $\cos\delta_U \to 1$, $\sin\delta_U \to \delta_U$, $\sin\delta_E \to \delta_E$, and $\delta_E \cdot \delta_U \to 0$, one may write $[BA]$ as[2]:

$$[BA] \cong \begin{bmatrix} 1 & -\delta_E & 0 \\ \delta_E & 1 & \delta_U \\ 0 & -\delta_U & 1 \end{bmatrix}$$

or

$$[BA] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & -\delta_E & 0 \\ \delta_E & 0 & \delta_U \\ 0 & -\delta_U & 0 \end{bmatrix} \tag{A-5}$$

[1] One must carefully distinguish between the angles (X, Y) and the moving two antenna axes (X, Y) and the space coordinates (x, y, z) for a target N under the reference axes (1), (2), and (3) in Fig. A-1.

[2] Equation (A-5) shows that $[AB] = [BA]$, which means that for small angles a reverse positioning of the axes will only reverse Eqs. (A-1) and (A-2).

Since the three direction cosines of the target position vector, ON, with respect to axes (1), (2) and (3), respectively are:

$$\cos\alpha = \cos Y \sin X \tag{A-6a}$$

$$\cos\beta = \sin Y \tag{A-6b}$$

$$\cos\gamma = \cos Y \cos X \tag{A-6c}$$

these may be used along with Eq. (A-5) in Eq. (A-4) to obtain three equations for $\cos\alpha'$, $\cos\beta'$ and $\cos\gamma'$ and solve for ΔX and ΔY as follows:

Using Eq. (A-6b) with Eq. (A-5) in Eq. (A-4) to obtain $\cos\beta'$ as

$$\sin Y' = \sin Y + \delta_E \cos Y \sin X + \delta_U \cos Y \cos X$$

Let

$$Y' = Y + \Delta Y$$

Hence,

$$\sin Y' = \sin Y \cos \Delta Y + \cos Y \sin \Delta Y$$

and for small ΔY

$$\sin Y + \cos Y \cdot \Delta Y = \sin Y + \delta_E \cos Y \sin X + \delta_U \cos Y \cos X$$

or

$$\Delta Y = \delta_E \sin X + \delta_U \cos X \tag{A-7}$$

Similarly from Eqs. (A-4), (A-5), and (A-6a), $\cos\alpha'$ is computed for small ΔX and ΔY,

$$(\cos Y - \sin Y' \cdot \Delta Y)$$

$$\times (\sin X + \cos X \cdot \Delta X) = \cos Y \sin X - \delta_E \sin Y$$

or

$$\cos Y \cos X \cdot \Delta X - \sin Y' \sin X \cdot \Delta Y = -\delta_E \sin Y \tag{A-8}$$

Also, from Eqs. (A-4), (A-5), and (A-6c), for small $(\Delta X, \Delta Y)$, $\cos\gamma'$ is written as

$$(\cos Y - \sin Y \cdot \Delta Y)$$

$$\times (\cos X - \sin X \cdot \Delta X) = \cos Y \cos X - \delta_U \sin Y$$

or

$$\cos Y \sin X \cdot \Delta X + \sin Y \cos X \cdot \Delta Y = \delta_U \sin Y \tag{A-9}$$

Multiplying Eq. (A-8) by $\cos X$, Eq. (A-9) by $\sin X$, adding and simplifying,

$$\Delta X = \tan Y [\delta_U \sin X - \delta_E \cos X] \tag{A-10}$$

Eq. (A-10) gives the correction ΔX in terms of δ_U and δ_E as shown in Eq. (A-1).

II. Antenna Angular Errors Due to Lack of Orthogonality

The two errors in X and Y angles, ΔX and ΔY, due to the lack of orthogonality of X, Y axes, δ_N, and between the positive RF and Y axes, δ_R, are derived. Both pairs of equations for ΔX and ΔY are used to solve for orthogonality errors, δ_N and δ_R.

First, the corrections, ΔX and ΔY, due to lack of orthogonality δ_N between the positive X and Y axes are given, in the following section, as

$$\tan \Delta X = \sin \delta_N \tan Y' \tag{A-11}$$

$$\tan \Delta Y = \frac{(1 - \cos \Delta X \cos \delta_N) \tan Y}{\tan^2 Y + \cos \delta_N \cos \Delta X} \tag{A-12}$$

for small δ_N, and $\tan \Delta Y < (\delta_N^2 + \delta_X^2)/2$.

Equations (A-11) and (A-12) reduce to:

$$\Delta X \approx \delta_N \tan Y' \tag{A-13}$$

$$\Delta Y \simeq 0 \tag{A-14}$$

See the description accompanying Figs. A-2 and A-3 for the explanation of symbols.

Second, the corrections, ΔX and ΔY, due to lack of orthogonality δ_R between the RF and Y axes are given from appendix section II-B as

$$\sin \Delta X = \frac{\sin \delta_R}{\cos Y} \tag{A-15}$$

$$\sin \Delta Y = \tan Y' (\cos \Delta Y - \cos \delta_R) \tag{A-16}$$

For small δ_R, these reduce to:

$$\Delta X \simeq \frac{\delta_R}{\cos Y'} \qquad \text{(A-17)}$$

$$\Delta Y \simeq 0 \qquad \text{(A-18a)}$$

where

$$\sin \Delta Y < \tan Y \cdot \delta_R^2$$

Since these corrections are very small, the total corrections to X and Y may be added and written as:

$$\Delta X_{total} \simeq \frac{\delta_N \sin Y' + \delta_R}{\cos Y'}$$

$$\Delta Y_{total} \simeq 0 \qquad \text{(A-18b)}$$

where

$$X \simeq X' - \Delta X$$

$$Y \simeq Y' \qquad \text{(A-18c)}$$

A. X and Y Axes Nonorthogonal

Figure A-2 is developed by first rotating X angle (from P_0 to M) about the antenna X axis, OP_2; then Y angle is rotated about the new Y axis. X angle below OP_1 is not shown. Thus OP_0 has been rotated to point N at the target in direction ON, about orthogonal axes. Secondly, consider the antenna Y axis OP_1 to be rotated so that it makes angle $(90 - \delta_N)^\circ$ with the antenna X axis OP_2. Now rotate X' angle from P_0 to P_0' about the antenna Y axis OQ. Thus OP_0 has been rotated to a point at the target in direction ON, about nonorthogonal axes.

The angle differences $\Delta X = X' - X$ and $\Delta Y = Y' - Y$ are the corrections to be subtracted from X' and Y' respectively to correct for this particular lack of orthogonality. Any Y axis error component about the X axis will be absorbed in the boresight error.

Consider the right spherical triangle $P_0'MN$ to obtain ΔX and ΔY.

$$\bar{\alpha} = (90 - \delta_N)^\circ, \quad \bar{\gamma} = 90^\circ$$

from spherical triangle relations:

$$\cos \bar{\alpha} = \tan \Delta X \cot Y'$$

Hence,

$$\sin \delta_N = \tan \Delta X \cot Y'$$

or

$$\tan \Delta X = \sin \delta_N \tan Y' \qquad \text{(A-19)}$$

Let

$$Y' = Y + \Delta Y,$$

Hence,

$$\tan (Y + \Delta Y) = \frac{\tan Y + \tan \Delta Y}{1 - \tan Y \tan \Delta Y}$$

by using Eq. (A-19)

$$\frac{\tan Y + \tan \Delta Y}{1 - \tan Y \tan \Delta Y} = \frac{\tan \Delta X}{\sin \delta_N} \qquad \text{(A-19a)}$$

Also, from the right triangle $P_0' MN$, spherical trigonometry gives

$$\sin \Delta X = \tan Y \tan \delta_N \qquad \text{(A-19b)}$$

substituting tan Y in Eq. (A-19a).

Hence,

$$\frac{\dfrac{\sin \Delta X}{\tan \delta_N} + \tan \Delta Y}{1 - \dfrac{\sin \Delta X \tan \Delta Y}{\tan \delta_N}} = \frac{\tan \Delta X}{\sin \delta_N}$$

or by cross multiplication:

$$\sin \Delta X \cos \delta_N + \tan \Delta Y \sin \delta_N = \tan \Delta X - \frac{\sin^2 \Delta X \tan \Delta Y}{\cos \Delta X \tan \delta_N}$$

and by rearranging tan ΔY terms:

$$\tan \Delta Y = \frac{\tan \Delta X - \sin \Delta X \cos \delta_N}{\sin \delta_N + \dfrac{\sin^2 \Delta X}{\cos \Delta X \tan \delta_N}}$$

Also, dividing by sin ΔX gives

$$\tan \Delta Y = \frac{\frac{1}{\cos \Delta X} - \cos \delta_N}{\frac{\tan \Delta X}{\tan \delta_N} + \frac{\sin \delta_N}{\sin \Delta X}}$$

substituting for tan δ_N and sin ΔX using Eq. (A-19b)

$$\tan \Delta Y = \frac{\frac{(1 - \cos \delta_N \cos \Delta X)}{\cos \Delta X}}{\frac{\tan \Delta X \tan Y}{\sin \Delta X} + \frac{\sin \delta_N}{\tan Y \tan \delta_N}}$$

Hence,

$$\tan \Delta Y = \frac{1 - \cos \delta_N \cdot \cos \Delta X}{\tan Y + \frac{\cos \delta_N \cos \Delta X}{\tan Y}} \qquad \text{(A-20)}$$

Substituting the Taylor series expansion:

$$\cos \delta_N \simeq 1 - \frac{\delta_N^2}{2}$$

$$\cos \Delta X \simeq 1 - \frac{\Delta X^2}{2}$$

and noting that the minimum value of the denominator in Eq. (A-20) is 2, the numerator becomes:

$$1 - \left(1 - \frac{\delta_N^2}{2}\right)\left(1 - \frac{\Delta X^2}{2}\right) \cong \frac{\delta_N^2}{2} + \frac{\delta_X^2}{2}$$

As tan $\Delta Y \approx \Delta Y$ for ΔY as small as indicated above, and since $\delta_N \leqslant \Delta X$ for $Y \geqslant 45°$

$$\Delta Y \leqslant \delta_N^2$$

or

$$\Delta Y \simeq 0, \quad Y \geqslant 45° \qquad \text{(A-21)}$$

In all cases $\Delta Y < \frac{\delta_N^2 + \Delta X^2}{2}$

Thus Eq. (A-19) becomes

$$\Delta X \simeq \delta_N \cdot \tan Y' \qquad \text{(A-22)}$$

B. Y and RF Axis Nonorthogonal

Figure A-3 similarly illustrates the angular corrections required by an angle $(90 - \delta_R)°$ between the Y axis of the antenna (OS_1) and RF(OT_1) axis as shown (OT_0), instead of 90°. The RF axis error component about the Y axis will be absorbed by the boresight error.

First, rotate the orthogonal RF axis angle X from T_1 to T_1' about the antenna X axis OS_3. Then rotate angle Y from T_1' to T_1'' about the new Y axis (not shown), which is angle X below OS_1. Second, rotate OT_0 by angle X' to OT_0' about the X axis. Note that OT_1 moved to OT_2. Then rotate angle Y' from T_0' to T_1'' about the new Y axis OS_2. Note that T_2 moves to T_2' under this Y' rotation. Arc $S_2T_1''T_2' = 90°$ and is part of a great circle and is $\angle S_2\ OT_2' = 90°$. Arc S_2S_3 is also part of a great circle whose plane is $\perp$ to $T_0'O$. Arc $S_2T_1''T_2' \perp$ arc $T_2T_2'S_3$ at T_2'.

In Fig. A-3, $\Delta X = (X' - X)$, so ΔX must be subtracted algebraically from X' to obtain the true X. $\Delta Y = (Y' - Y)$, so ΔY must be algebraically subtracted from Y' to obtain the true Y. Spherical triangle $S_2S_3T_1''$ of Fig. A-3 provides solutions for ΔX and ΔY. Since

$$\cos (90 - \delta_R) = \cos 90 \cos (90 - Y) + \sin (90 - Y) \cos (90 - \Delta X)$$

Hence,

$$\sin \delta_R = \cos Y \sin \Delta X$$

or

$$\sin \Delta X = \sin \delta_R / \cos Y \qquad \text{(A-23)}$$

Also, since

$$\cos (90 - Y) = \cos (90 - \delta_R) \cos 90 + \sin (90 - \delta_R) \sin 90 \cos (90 - Y')$$

then

$$\sin Y = \cos \delta_R \sin Y' \qquad \text{(A-24)}$$

Substituting

$$Y = Y' - \Delta Y,$$

Equation (A-24) gives:

$$\sin Y' \cos \Delta Y - \cos Y' \sin \Delta Y = \cos \delta_R \sin Y'$$

or

$$\sin \Delta Y = \frac{-\cos \delta_R \sin Y' + \sin Y' \cos \Delta Y}{\cos Y'}$$

or

$$\sin \Delta Y = \tan Y' (\cos \Delta Y - \cos \delta_R) \qquad \text{(A-25)}$$

Similarly by substituting $Y' = (Y + \Delta Y)$ into Eq. (A-24) gives

$$\sin Y = \cos \delta_R \sin Y \cos \Delta Y + \cos \delta_R \cos Y \sin \Delta Y$$

or

$$\sin \Delta Y = \frac{\sin Y - \cos \delta_R \sin Y \cos \Delta Y}{\cos \delta_R \cos Y}$$

or

$$\sin \Delta Y = \left(\frac{1}{\cos \delta_R} - \cos \Delta Y \right) \tan Y \qquad \text{(A-26)}$$

Since for $90° > \begin{pmatrix} Y \\ \text{and} \\ Y' \end{pmatrix} > 0, \begin{pmatrix} \tan Y \\ \text{and} \\ \tan Y' \end{pmatrix} > 0$

and since

$$\left(\frac{1}{\cos \delta_R} - \cos \Delta Y \right) > 0 \text{ due to } \frac{1}{\cos \delta_R} > \cos \Delta Y,$$

$\sin \Delta Y > 0$ from Eq. (A-26), which makes $\cos \Delta Y > \cos \delta_R$

$$\left.\begin{array}{l} \cos \Delta Y > \cos \delta_R \\ \delta_R > \Delta Y \end{array}\right\} \qquad \text{(A-27)}$$

From Eq. (A-26):

$$\sin \Delta Y < \tan Y \left[1 + \frac{\delta_R^2}{2} - \left(1 - \frac{\Delta Y^2}{2} \right) \right]$$

From Eq. (A-27)

$$\sin \Delta Y < \tan Y (\delta_R^2) \qquad \text{(A-28)}$$

or

$$\Delta Y \simeq 0$$

Equation (A-23) then becomes Eq. (A–17):

$$\Delta X \simeq \frac{\delta_R}{\cos Y'} \qquad \text{(A-29)}$$

III. Partial Derivatives of *X* and *Y* Angles and Associated Error Terms

Errors in X and Y angles due to an error in a parameter P may be expressed by retaining only the linear term of a Taylor series expansion of the parameter so that

$$P = P_e + \Delta P$$

where P_e is the estimated value of P, and ΔP is the error in the estimate of P. Then the error in X (and similarly in Y) resulting from errors in estimating n parameters is,

$$\Delta X = \sum_{i=1}^{n} \left(\frac{\partial X}{\partial P} \right)_i \Delta P_i, \quad \Delta Y = \sum_{i=1}^{n} \left(\frac{\partial Y}{\partial P} \right)_i \Delta P_i$$

IV. Correction Due to Errors in Station Location

Errors in station latitude and longitude $\Delta\phi$ and $\Delta\lambda$ cause errors in X and Y angles as shown in Eqs. (A-30) and (A-31) below:

$$\Delta X = \frac{\partial X}{\partial \phi} \Delta\phi + \frac{\partial X}{\partial \lambda} \Delta\lambda \qquad \text{(A-30)}$$

$$\Delta Y = \frac{\partial Y}{\partial \phi} \Delta\phi + \frac{\partial X}{\partial \lambda} \Delta\lambda \qquad \text{(A-31)}$$

The partial derivatives are shown below. X and Y angles of a target can be related to the local hour angle t and declination δ of the target by

$$X = \sin^{-1}\left(\frac{-\cos\delta\sin t}{\cos Y}\right) \quad \text{(A-32)}$$

$$Y = \sin^{-1}(\sin\delta\cos\phi - \cos\delta\sin\phi\cos t) \quad \text{(A-33)}$$

Local hour angle t is a function of constants and of station longitude, so the partial derivatives with respect to t are equivalent to partial derivative with respect to λ.

Taking partial derivatives of Eqs. (A-32) and (A-33) with respect to ϕ and λ (or t) yields the four partial derivatives required by Eqs. (A-30) and (A-31).

The results are listed as:

$$\left(\frac{\partial Y}{\partial\lambda}\right)_\phi = \frac{\sin\phi\cos\delta\sin t}{\sqrt{1-(\sin\delta\cos\phi - \sin\phi\cos\delta\cos t)^2}} \quad \text{(A-34)}$$

$$\left(\frac{\partial X}{\partial\lambda}\right)_\phi = \frac{-\cos Y\cos\delta\cos t - \cos\delta\sin t\sin Y\left(\frac{\partial Y}{\partial\lambda}\right)}{\cos^2 Y\sqrt{1-\left(\frac{\cos\delta\sin t}{\cos Y}\right)^2}} \quad \text{(A-35)}$$

$$\left(\frac{\partial Y}{\partial\phi}\right)_\lambda = \frac{-\sin\delta\sin\phi - \cos\phi\cos\delta\cos t}{\sqrt{1-(\sin\delta\cos\phi - \sin\phi\cos\delta\cos t)^2}} \quad \text{(A-36)}$$

$$\left(\frac{\partial X}{\partial\phi}\right)_\lambda = \frac{-\cos\delta\sin t\sin Y\left(\frac{\partial Y}{\partial\phi}\right)}{\cos^2 Y\sqrt{1-\left(\frac{\cos\delta\sin t}{\cos Y}\right)^2}} \quad \text{(A-37)}$$

V. Errors Due to Structural Deflection

As described earlier, X and Y deflection errors are assumed to be maximum at a near-horizon-reference $\bar{\phi}$ and minimum at an elevation angle $(h - \bar{\phi}) = 90°$, where h is the antenna elevation angle. Deflection errors in an elevation plane are transformed into X and Y errors by using partial derivatives with respect to h.

The deflection errors in X and Y angles are

$$\Delta X = C_1\cos(h - C_2)\left(\frac{\partial X}{\partial h}\right)_\psi \quad \text{(A-38)}$$

$$\Delta Y = C_4\cos(h - C_5)\left(\frac{\partial Y}{\partial h}\right)_\psi \quad \text{(A-39)}$$

The partial derivatives are derived below utilizing the X-Y to azimuth (ψ), elevation (h) transformations. The partials required by Eqs. (A-38) and (A-39) are given in Eqs. (A-47) and (A-48), where the derivation is made as follows:

$$\sin h = \cos Y\cos X \quad \text{(A-40)}$$

$$\cos h\sin\psi = \cos Y\sin X \quad \text{(A-41)}$$

$$\sin Y = \cos h\cos\psi \quad \text{(A-42)}$$

Partial derivative with respect to h of (A-42) gives:

$$\cos Y\left(\frac{\partial Y}{\partial h}\right)_\psi = -\sin h\cos\psi$$

And substitution of $\cos\psi$ from (A-42) delivers

$$\left(\frac{\partial Y}{\partial h}\right)_\psi = \frac{-\sin h}{\cos Y}\cdot\frac{\sin Y}{\cos h} = -\tan h\tan Y \quad \text{(A-43)}$$

Division of (A-41) by (A-40) provides,

$$\frac{\cos Y\sin X}{\cos Y\cos X} = \frac{\cos h\sin\psi}{\sin h}$$

or

$$\tan X = \cot h\sin\psi \quad \text{(A-44)}$$

Differentiating (A-44) with respect to h,

$$\sec^2 X\left(\frac{\partial X}{\partial h}\right)_\psi = -\csc^2 h\sin\psi \quad \text{(A-45)}$$

Use of Eqs. (A-40) and (A-41) simplifies (A-45) to

$$\left(\frac{\partial X}{\partial h}\right)_\psi = \frac{-\cos Y\sin X\cos^2 X}{\cos h\sin^2 h} = \frac{-\cos Y\sin X}{\cos h\cos^2 Y}$$

$$= \frac{-\sin X}{\cos h\cos Y} \quad \text{(A-46)}$$

Finally from (A-43) and (A-46)

$$\left(\frac{\partial X}{\partial h}\right)_{\psi=const} = -\frac{\sin X}{\cos h\cos Y} \quad \text{(A-47)}$$

$$\left(\frac{\partial Y}{\partial h}\right)_{\psi=const} = -\tan h \tan Y \qquad \text{(A-48)}$$

VI. Boresight Shift with Polarization Rotation

A common characteristic of the circularly polarized monopulse tracking antenna is that a boresight shift occurs as the receive signal polarization vector is rotated with respect to the receiving antenna feed. The greater the axial ratio of the received polarized signal, the greater the magnitude of the boresight shift. The boresight shift varies in a systematic manner with the rotation of the polarization vector of the incident electromagnetic waves. In general, a relative rotation of the polarization vectors of 180 degrees will result in the boresight of the monopulse antenna traversing an ellipse (in the orthogonal angular coordinates of the antenna) centered on the nominal boresight. In general, the orientation of the ellipse relative to the orthogonal angular coordinate of the antenna is arbitrary.

In general, all aircraft tracks are commonly performed with a linear (vertical) polarized transmit signal and an elliptically polarized receiving antenna. The motion of the X-Y mount antenna is such that, throughout tracking, the orientation of the tracking feed is rotated about the line-of-sight axis. For a typical track, this can result in a significant shifting of the system boresight throughout the track. Thus it is necessary that the boresight shift with polarization rotation error term be included in the reduction of tracks in order to determine the proper values for the other error model terms.

Referencing the feed system space angles, the Y and cross-Y components of the boresight shift are given by the generalized parametric equations of an ellipse:

$$ERROR_{CY} = A_9 \sin 2\theta \qquad \text{(A-49)}$$

$$\begin{aligned} ERROR_Y &= M \sin(2\theta + N) \\ &= M(\sin 2\theta \cos N + \cos 2\theta \sin N) \\ &= A_{10} \sin 2\theta + A_{11} \cos 2\theta \end{aligned} \qquad \text{(A-50)}$$

where

$A_{10} = M \cos N$,

$A_{11} = M \sin N$,

θ is the angle between a reference line on the feed which is vertical (when the XY mount faces the east or south horizon depending on the antenna design and a constantly vertical line, representing the vertical transmit polarization).

N is the angle between the radial lines which pass through the tangent points of the ellipse and the lines $ERROR_{CY} = A_9$ and $ERROR_Y = M$, respectively (refer to Fig. A-4).

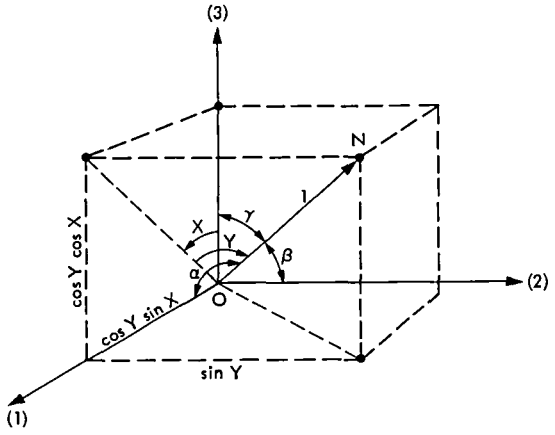
$$\theta = \tan^{-1}\left(\frac{\cos X \sin Y}{\sin X}\right)$$

and since the regression equations are solved in X angle (antenna shaft angle) rather than cross-Y (space angle), Eqs. (A-49) and (A-50) become:

$$ERROR_X = \frac{ERROR_{CY}}{\cos Y} = \frac{A_9 \sin 2\theta}{\cos Y} \qquad \text{(A-51)}$$

$$ERROR_Y = A_{10} \sin 2\theta + A_{11} \cos 2\theta \qquad \text{(A-52)}$$

Equations (A-51) and (A-52) were incorporated into the total system angular error model for reduction of the system test data. Note that after regression estimates of A_9, A_{10}, and A_{11} are obtained, the maximum boresight shift with polarization of the antenna can be determined.



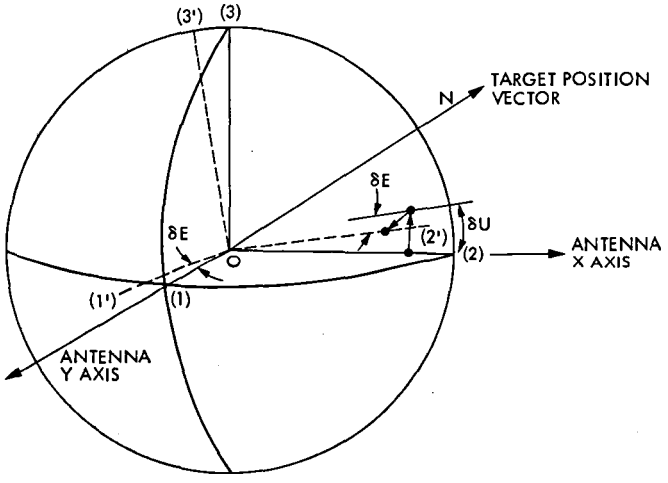
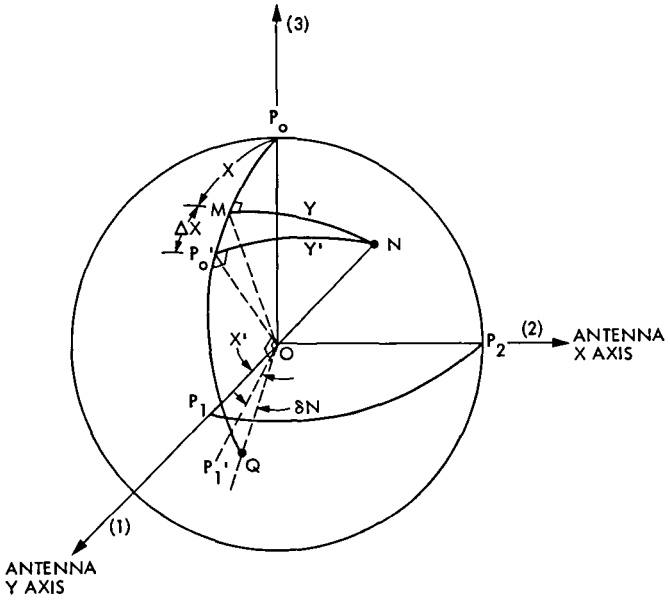


Fig. A-1. Angle error derivations



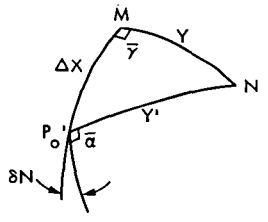


Fig. A-2. Orthogonality error derivations (1)

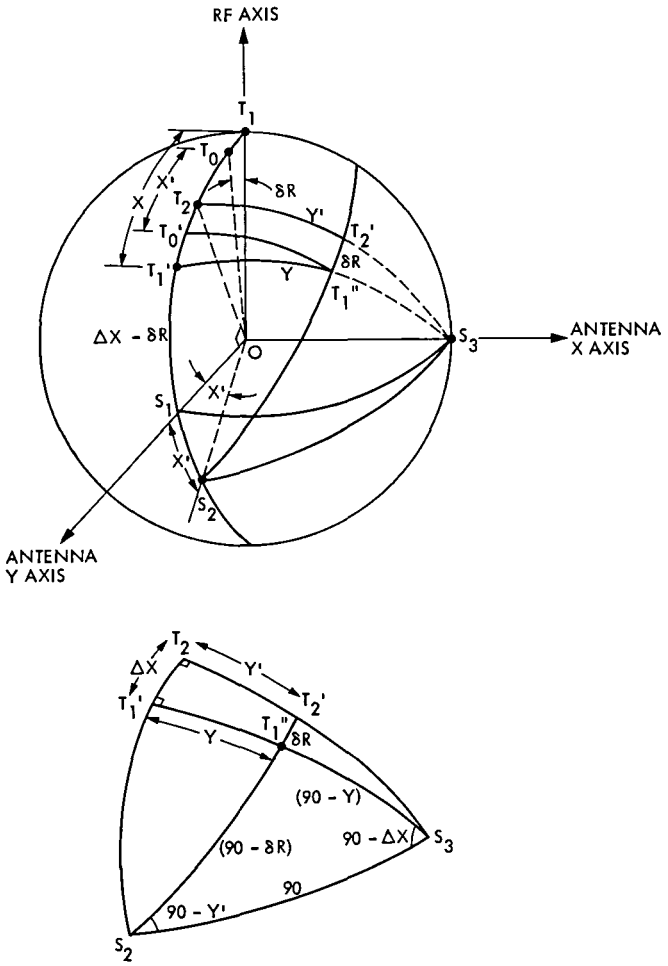


Fig. A-3. Orthogonality error derivations (2)

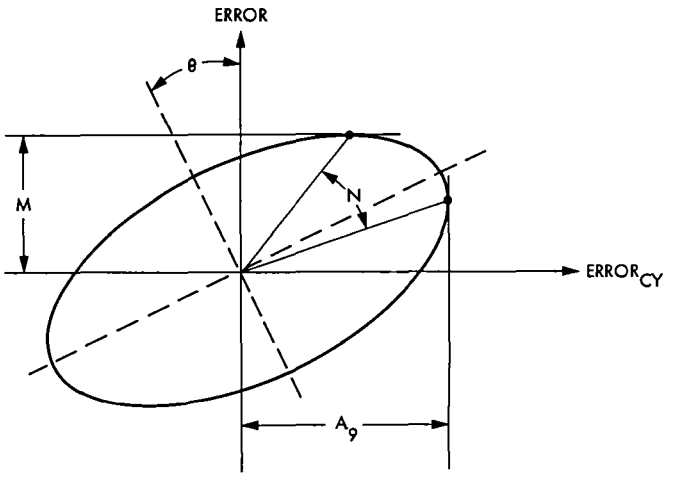


Fig. A-4. Geometrical relationships for boresight shift with polarization model

Root Locus Algorithms for Programmable Pocket Calculators

E. R. Wechsler
Communications Systems Research Section

Two algorithms are described which allow the plotting of individual points on a root locus diagram with or without time delay. The development was performed during the design of a continuous phase shifter used in the Baseband Antenna Combiner for the Deep Space Network (DSN). The algorithms, which are expected to be useful in similar DSN efforts, are simple enough to be implemented on a programmable pocket calculator. The coordinates of the open-loop zeros and poles, the gain constant K, and the time delay T are the data inputs.

I. Introduction

The root locus method (Ref. 1) allows the designer to obtain the poles of a closed-loop linear system when the zeros and poles of the open loop are known. The equation describing the root locus is

$$K \frac{\prod_{i=1}^{z} (s - s_i)}{\prod_{i=z+1}^{z+p} (s - s_i)} e^{-Ts} = -1 \qquad (1)$$

where

$s = \sigma + j\omega$

$s_i = \sigma_i + j\omega_i$ { are zeros for $1 \leqslant i \leqslant z$; are poles for $z < i \leqslant z + p$

$T \geqslant 0$ is the time delay

K is a real parameter (usually a gain constant)

There are basically three methods for plotting root loci by digital computation:

(1) The s plane is scanned until points of the locus are found with the desired accuracy (Ref. 2). This method requires lots of computation and of memory.

(2) The plotting starts from a pole and advances along the locus in fixed increments. If any parameter is modified, the search has to start again because the locus changes its shape (Refs. 3, 4).

(3) The plotting is performed by solving the polynomial:

$$Q(s) + KP(s) = 0$$

This approach was implemented on a pocket calculator (Ref. 5). It does not apply to systems with time delay and it requires a preliminary expansion of the polynomial by hand or with an additional program.

The root locus algorithms described here were developed during the design of a continuous phase shifter used in the Baseband Antenna Combiner of the DSN. The algorithms

should be helpful to engineers involved in the design of linear feedback systems.

II. Calculations

Using the complex Ln function we get from (1):

$$\mathrm{Ln}\left[K\frac{\prod_{i=1}^{z}(s-s_i)}{\prod_{i=z+1}^{z+p}(s-s_i)}e^{-Ts}\right]-\mathrm{Ln}\,(-1)=0 \qquad (2)$$

which can be expanded to:

$$\ln|K|+\sum_{i=1}^{z+p}\delta_i \ln|s-s_i|-\sigma T+j\left[\sum_{i=1}^{z+p}\delta_i \arg(s-s_i)-\omega T-n\pi\right]=0 \qquad (3)$$

$$n=\begin{cases}\pm1,\pm3,\pm5,\cdots \text{ for } K>0\\ 0,\pm2,\pm4,\cdots \text{ for } K<0\end{cases}$$

$$\delta_i=\begin{cases}1 \text{ for zeros } 1\leqslant i\leqslant z\\ -1 \text{ for poles } z<i\leqslant z+p\end{cases}$$

Two real functions are defined, based on the real and the imaginary parts of (3):

$$R(\sigma,\omega)=\ln|K|+\sum_{i=1}^{z+p}\delta_i \ln\left[(\sigma-\sigma_i)^2+(\omega-\omega_i)^2\right]^{1/2}-\sigma T \qquad (4a)$$

$$Q(\sigma,\omega)=\sum_{i=1}^{z+p}\delta_i \arg\left[(\sigma-\sigma_i)+j(\omega-\omega_i)\right]-\omega T-n\pi \qquad (4b)$$

n such that $-\pi<Q\leqslant\pi$

A point on the root locus will satisfy the equations:

$$R(\sigma,\omega)=0 \quad \text{and} \quad Q(\sigma,\omega)=0 \qquad (5)$$

Two different methods were used for solving Eqs. (5):

The Newton-Raphson iteration (Ref. 6) starts from a point of coordinates σ and ω and produces the correction terms $\Delta\sigma$ and $\Delta\omega$ from the following equations:

$$R(\sigma,\omega)+R_\sigma\Delta\sigma+R_\omega\Delta\omega=0 \qquad (6a)$$

$$Q(\sigma,\omega)+Q_\sigma\Delta\sigma+Q_\omega\Delta\omega=0 \qquad (6b)$$

Since $R(\sigma,\omega)$ and $Q(\sigma,\omega)$ are the real and imaginary parts of the analytic function (2) they satisfy the Cauchy-Riemann conditions

$$Q_\sigma=-R_\omega \qquad (7a)$$

$$Q_\omega=R_\sigma \qquad (7b)$$

Substituting (7) in (6) and solving the system we get:

$$\Delta\sigma=\frac{-R\cdot R_\sigma+Q\cdot R_\omega}{R_\sigma^2+R_\omega^2} \qquad (8a)$$

$$\Delta\omega=\frac{-R\cdot R_\omega-Q\cdot R_\sigma}{R_\sigma^2+R_\omega^2} \qquad (8b)$$

R_σ and R_ω are derived from (4a):

$$R_\sigma=\sum_{i=1}^{z+p}\delta_i\frac{\sigma-\sigma_i}{(\sigma-\sigma_i)^2+(\omega-\omega_i)^2}-T \qquad (9a)$$

$$R_\omega=\sum_{i=1}^{z+p}\delta_i\frac{\omega-\omega_i}{(\sigma-\sigma_i)^2+(\omega-\omega_i)^2} \qquad (9b)$$

The coordinates of the new iteration are:

$$\sigma'=\sigma+\Delta\sigma \quad \text{and} \quad \omega'=\omega+\Delta\omega \qquad (10)$$

The iteration process continues until:

$$\Delta\sigma^2+\Delta\omega^2\leqslant P^2$$

where P is the desired precision.

This algorithm converges very fast but it presents difficulties when the starting point is far from the solution because it can diverge and give errors or jump to a solution on a branch other than the one of interest.

A gradient method (Ref. 7) was developed which does not have the disadvantages of the previous algorithm but is slower to converge. Consider the gradients of $R(\sigma, \omega)$ and $Q(\sigma, \omega)$:

$$\text{grad } R = \mathbf{u}R_\sigma + \mathbf{v}R_\omega \tag{11a}$$

$$\text{grad } Q = \mathbf{u}Q_\sigma + \mathbf{v}Q_\omega \tag{11b}$$

u and **v** are unit vectors along the respective σ and ω axis.

Substituting (7) in (11b):

$$\text{grad } Q = -\mathbf{u}R_\omega + \mathbf{v}R_\sigma \tag{12}$$

We notice that:

$$|\text{grad } R| = |\text{grad } Q| \tag{13a}$$

$$\arg(\text{grad } Q) - \arg(\text{grad } R) = \frac{\pi}{2} \tag{13b}$$

Because the two gradients are orthogonal, moving along grad R will produce a maximum rate of change for R and a zero rate of change for Q and vice versa. In order to make R decrease in absolute value at the fastest rate we have to move in the direction pointed by the vector:

$$\mathbf{a} = -\text{Sgn}(R) \cdot \text{grad } R \qquad \text{Sgn}(R) = 1 \text{ for } R \geqslant 0 \qquad \text{Sgn}(R) = -1 \text{ for } R < 0$$

Likewise for Q we have to move along:

$$\mathbf{b} = -\text{Sgn}(Q) \cdot \text{grad } Q$$

Therefore moving along the search vector $\mathbf{c} = \mathbf{a} + \mathbf{b}$ will reduce both R and Q in absolute value.

$$\mathbf{c} = -\text{Sgn}(R) \cdot \text{grad } R - \text{Sgn}(Q) \cdot \text{grad } Q \tag{14}$$

Depending on the polarities of R and Q there are four possible values for the angle θ between **c** and grad R (Fig. 1).

$$\theta = \arg(\mathbf{c}) - \arg(\text{grad } R) \tag{15}$$

$$\left.\begin{array}{ll} R \geqslant 0 \;;\; Q \geqslant 0 & \theta = -135^\circ \\ R \geqslant 0 \;;\; Q < 0 & \theta = 135^\circ \\ R < 0 \;;\; Q \geqslant 0 & \theta = -45^\circ \\ R < 0 \;;\; Q < 0 & \theta = 45^\circ \end{array}\right\} \tag{16}$$

We start from a point of coordinates σ and ω and we calculate R, Q, R_σ, R_ω from (4a), (4b), (9a), (9b) respectively.

We find arg (grad R) from R_σ and R_ω using the rectangular to polar conversion. Knowing the polarities of R and Q we find θ from table (16). We get the direction of search vector **c** from (15):

$$\arg(\mathbf{c}) = \arg(\text{grad } R) + \theta$$

During one pass, the magnitude of the step is a constant D. The recommended value for D during the first pass is a few percent of full scale.

The search vector has the modulus D and the argument arg (**c**). Using the polar to rectangular conversion we get the coordinate corrections $\Delta\sigma$ and $\Delta\omega$. The search moves now to a new point of coordinates:

$$\sigma' = \sigma + \Delta\sigma \qquad \omega' = \omega + \Delta\omega$$

The search pass will consist initially of steps moving at a constant angle θ relative to the local grad R. As soon as the search crosses one of the curves $R = 0$ or $Q = 0$, it starts a zigzag pattern along the crossed curve because of the change in polarity of the corresponding function (Fig. 2). When the search reaches the vicinity of the solution ($R = 0$; $Q = 0$) the zigzag pattern is interrupted and a 180-deg change of direction occurs.

There are only two possible cases as shown in Fig. 2. The history of the last three steps is enough for detecting this 180-deg change in direction between the last step and any of the previous two steps.

This part of the algorithm is implemented through the use of a three-register stack into which the angles θ are fed. After each step we look for:

$$|\theta_N - \theta_{N-1}| = 180^\circ \quad \text{or} \quad |\theta_N - \theta_{N-2}| = 180^\circ$$

If the answer is true the step size D is reduced and the search process continues with a new pass. The search ends when the magnitude of the increment D becomes smaller than the required precision.

III. Conclusion

Both algorithms were implemented on a Casio FX-602P programmable pocket calculator. The first program uses 36 memory registers and 281 program steps, while the second one uses 41 memory registers and 351 program steps. They allow the design of systems with up to eight singularities ($z + p = 8$).

References

1. D'Azzo, J. J., and Houpis, C. H., *Linear Control System Analysis and Design*, McGraw-Hill, pp. 202-241, 1975.
2. Krall, A. M., and Fornaro, R., "An Algorithm for Generating Root Locus Diagrams," *Communications of the ACM*, Vol. 10, No. 3, pp. 186-188.
3. Williamson, S. E., "Accurate Root Locus Plotting Including the Effects of Pure Time Delay," *Proc. IEE (GB)*, Vol. 116, No. 7, pp. 1269-71.
4. Ash, R. H., and Ash, G. R., "Numerical Computation of Root Loci Using the Newton-Raphson Technique," *IEEE Trans.*, AC-13, pp. 576-58, 1968.
5. Harden, R. C., and Simons, F. O., "Root Locus Algorithms and Routines Adapted to Hand-Held HP-67 Computers," 12th Annual Southeastern Symposium on System Theory, Publ, *IEEE*, pp. 285-9, 1980.
6. Scarborough, J. B., *Numerical Mathematical Analysis*, 2nd ed., Oxford University Press, pp. 203-204, 1955.
7. Korn, G. A., and Korn, T. M., *Electronic Analog and Hybrid Computers*, 2nd ed., McGraw-Hill, pp. 328-335, 1972.

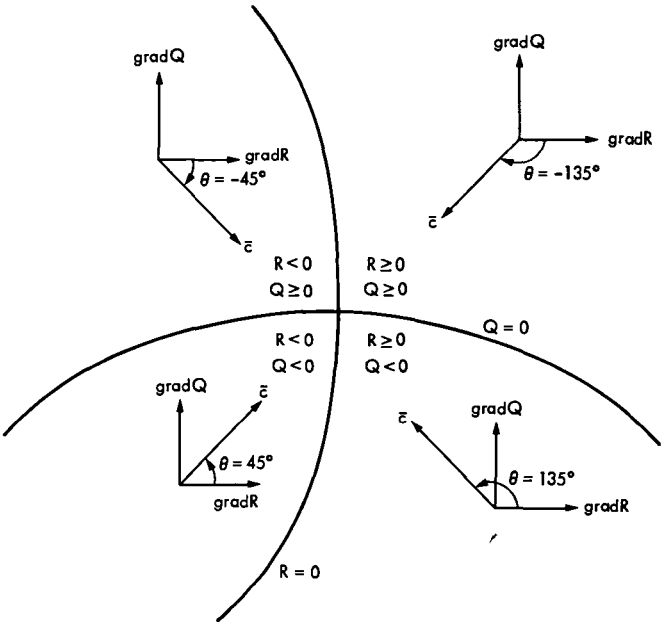


Fig. 1. Selection of search vector for the gradient method

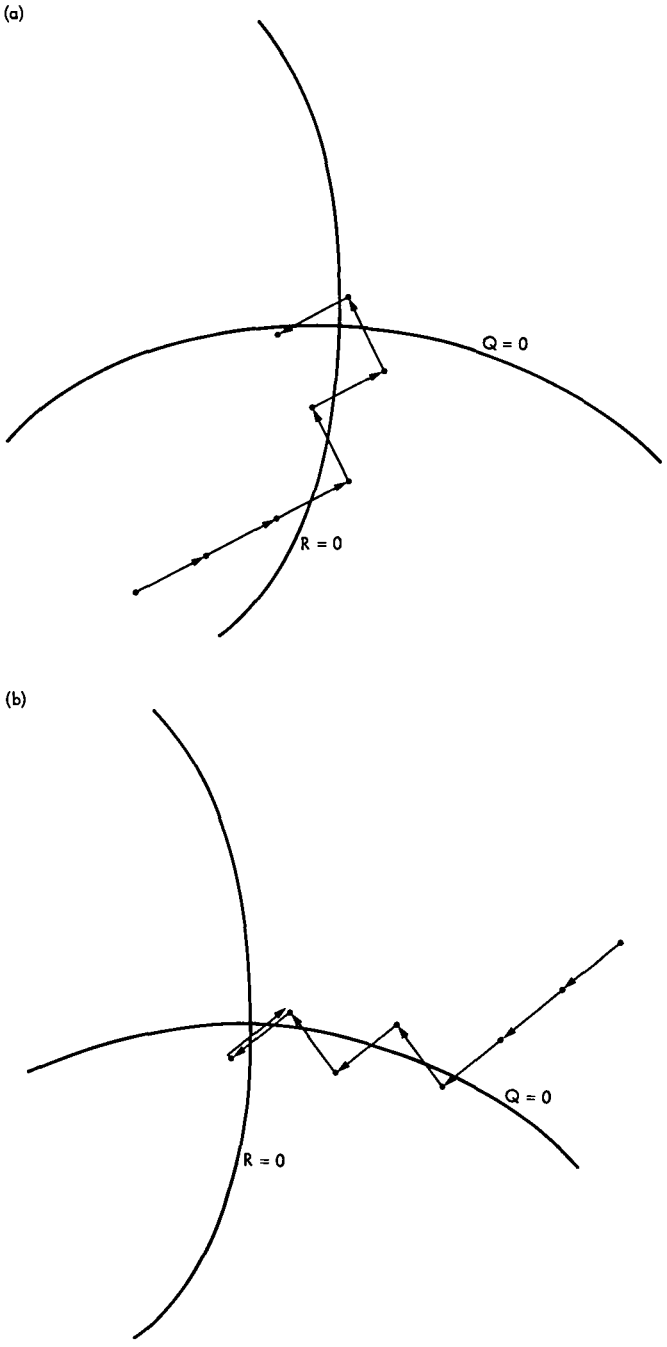


Fig. 2. Gradient method search pattern – 180-deg change of direction within the last three steps

A Covariance Analysis for the Determination of Baselines Observing GPS Satellites

S. C. Wu
Tracking Systems and Applications Section

Results of covariance analyses are presented for the determination of baseline vectors of length from 150 m to 4000 km. The data are doubly differenced ranges derived from the signals transmitted by the satellites of the Global Positioning System. The modeling of various error sources is described. The results indicate the low-cost baseline determination system is capable of sub-decimeter accuracy for baselines up to 1000 km in length.

I. Introduction

The application of Very Long Baseline Interferometry (VLBI) to the determination of baseline vectors has been investigated for over a decade (Refs. 1-5). The results of ARIES (Astronomical Radio Interferometric Earth Surveying) experiments have demonstrated an accuracy better than 10 cm for baselines up to a few hundred kilometers in length (Refs. 3, 4). The *lengths* of intercontinental baselines (~10,000 km) between the Deep Space Network stations in California, Spain, and Australia have also been determined to 10 cm accuracy using a more sophisticated VLBI system (Ref. 5).

The high precision in baseline determination is a result of observing extragalactic radio sources (quasars) with time-invariant angular positions which are accurately known. Utilizing the faint signals of quasars requires low-noise receivers coupled with high-precision clock, high-efficiency antennas of moderately large aperture, and a high-density data acquisition facility. The data from each end of the baselines, typically of the order of 10^{10} bits, are brought together for processing. Therefore, the operating cost of a VLBI system is quite high. For frequent monitoring of crustal motion in a seismically active area, a low-cost alternative is highly desirable.

The Global Positioning System (GPS), to be fully in operation after 1988, will consist of 18 navigation satellites evenly distributed in 6 orbit planes at an altitude of ~20,000 km. The use of signals transmitted by these GPS satellites in place of quasar signals for baseline determination has been proposed and studied by several researchers (Refs. 6, 7). The advantage of the GPS signals over quasar signals is that they are strong and encoded. This enables the use of a low-cost receiver with a small antenna and allows on-site preprocessing to reduce the data volume before being brought together for further processing. Such a system will provide low-cost operation and may become a strong candidate to supplement the quasar-based VLBI system to provide frequent baseline motion monitoring.

Pseudorange and range rate information can be easily extracted from the GPS signals without the knowledge of the signal codes by a system called SERIES (Satellite Emission

Range Inferred Earth Surveying) proposed by MacDoran (Ref. 7). This SERIES system is currently under development at the Jet Propulsion Laboratory and a series of tests on baseline determination capability is underway. This report provides a covariance analysis to estimate the accuracies with which baselines of different lengths can be determined by the SERIES approach.

Recently, the use of SERIES in the orbit determination of a low-altitude Earth satellite (e.g., TOPEX) has been investigated (Refs. 8, 9). Preliminary studies indicate a sub-decimeter altitude determination capability. While an in-flight proof-of-concept demonstration is impossible without actually implementing the whole system, including a dedicated satellite carrying a SERIES receiver, consistency between the results of the current analysis and of the baseline determination tests will provide increased confidence in predicted performance of the proposed SERIES satellite tracking system.

II. Data Types

The GPS satellites emit signals at two L-band frequencies, 1575.42 MHz (L_1) and 1227.60 MHz (L_2), that can be used to make one-way (pseudo-) range and doppler measurements between the GPS satellites and a user receiver. A detailed description of GPS signal structure can be found in Ref. 10. When the range or doppler measurements from a GPS satellite to the two ends of a baseline are brought together and differenced, the data are analogous to VLBI data. Differenced GPS range (or doppler) reduces the sensitivity to the GPS ephemeris errors and eliminates the GPS instrument (including clock) errors. However, the clock discrepancy between the two ends of the baseline remains as an error source to the baseline determination. The clock discrepancy will be the same no matter which satellite originates the signals. Then, when simultaneous differenced range (or doppler) measurements from two GPS satellites are further differenced, the clock error will cancel. These new data types are called doubly differenced GPS range (DDGR) and doubly differenced GPS doppler (DDGD), respectively (Ref. 9). Since the angular separation between the GPS satellites can be very large, the geometrical strength is not expected to be significantly reduced from that of singly differenced data.

In the following analysis the DDGR data taken at 5-minute intervals are used. The DDGD data will be only briefly studied.

III. Geometry

Although 18 GPS satellites will ultimately be in operation, only 6 satellites are currently in orbit, of which five are under operating conditions. Periodically, all five satellites can be in view nearly simultaneously from a station on the continental U.S. The view periods of these satellites on May 9, 1982, from a station in the state of California are shown in Fig. 1. A cutoff elevation angle of 20-deg was assumed. These view periods will be used in the following analysis. Other satellite parameters are summarized in Table 1.

Baselines to be studied are of lengths from 150 m to 4000 km. These are also listed in Table 1. A short baseline experiment is useful in providing a measure of how the instrument will perform, because the effects of atmospheric delays, of station location error, and of GPS ephemeris error will cancel nearly perfectly between the two ends of such a baseline.

IV. Error Models

The error models are summarized in Table 2. Since dual-frequency (L_1 and L_2) observations are available, ionospheric delay is assumed to be totally removed and is not considered here. Clock errors are also totally removed by the double difference inherent in the DDGR data.

The ephemeris error of each GPS satellite is represented by a full covariance matrix generated by a 9-hour orbit fit such that the RSS position error in three orthogonal directions remains nearly at 10 meters. A diagonal covariance matrix would cause the in-track position error to increase rapidly with time, as depicted by Fig. 2, and thus would result in an unrealistically large baseline determination error.

For the tropospheric error, we assume a 5-cm zenith calibration using surface measurements (Ref. 11) or a seasonal model (Ref. 12). Azimuth homogeneity is assumed so that the error is scaled only by an elevation factor. Over short baselines the effects of the troposphere between the two ends are highly correlated. The difference in zenith tropospheric delay between the two ends of a baseline of length B can be represented by a function proportional to $B^{0.7}$ (Ref. 13). Even though this relationship was derived from observations with B ranging from 1 to 20 km, the trend seems to suggest a linear extrapolation to $B \sim 80$ km. Within this range, a smaller error will result when a single zenith calibration is applied to both ends of the baseline. The proportionality factor 0.2 is twice as large as given in the reference to account for lower altitudes than the 2 km at the VLA site in New Mexico, where the data of the reference were acquired. For $B > 100$ km, the correlation breaks down to a level that independent calibration at each end of the baseline becomes a better choice.

With VLBI observation of quasars at an "infinite distance," the location error of one end of the baseline will translate

one-to-one into the location of the other end, resulting in no net *baseline* error. With DDGR measurements over long baselines, the finite distances of the GPS satellites will translate the location error between the two ends of the baseline less perfectly and a net baseline error will result. The assumed 1-m error in each component of the reference station (first of the pair) is relative to the "center of the Earth" defined by the GPS satellite orbits or by the station network determining such orbits.

The effects of polar motion and UT1 errors will rotate only the baseline orientation. These effects are independent of the observation scheme and can be easily determined (Ref. 14). These errors are not included in the current analysis.

V. Results of Covariance Analysis

Covariance analysis was performed using simulated DDGR data. The baseline vector determination errors as a function of baseline length are shown in Figs. 3-6, respectively, for the four error sources. The east, north and vertical components of these errors are relative to the second station of each baseline. All four error sources have their largest effects in the vertical component. The component along the north-south direction is least affected by data noise and GPS ephemeris error. The component along the transverse horizontal direction is least affected by tropospheric error and location error of the reference stations.

The effects of GPS emphemeris error and of the reference station location error increase with baseline length. For shorter baselines, the correlation of troposphere between the two stations decreases as baseline length, thus increasing its effects. For longer baselines, these effects level off at about 20 cm, still mostly in the vertical component. The effects of data noise remain nearly constant for most baseline lengths, with a slight increase for the longest baseline (Haystack-OVRO) due mainly to shorter common view periods of all GPS satellites from the two ends of the baseline.

The three components of the total (root-sum-square) baseline vector determination error are shown in Fig. 7, again as a function of baseline length. The shorter baselines are dominated by data noise. Tropospheric error dominates the baselines of medium length. For baselines longer than ~600 km, GPS ephemeris error becomes the dominating error source. While decimeter accuracy for the vertical component can be retained only for baselines shorter than 60 km, such accuracy can be retained for the horizontal components for baselines up to about 500 km in length.

An analysis was also performed for the DSS 14-DSS 13 baseline using DDGD data. Such data have the advantage over DDGR data in that they are more accurate[1], less affected by multipath effects and can be acquired with simpler receivers. However, it is well known that the information content of doppler data is less than the corresponding range data. This is confirmed in Fig. 8 by the larger data noise effects. Note that the 0.05-mm/s DDGD data noise at a 5-minute count time corresponds to 1 cm DDGR error, which is a factor of 10 smaller than the assumed data noise for the previous analysis using DDGR data. Such low DDGD data noise may be achievable from the GPS RF carriers.

The effects of tropospheric error are also smaller when using DDGD data. Even though the total baseline vector errors are larger than the corresponding DDGR solutions, these are not significant for longer baselines because the dominating effects of GPS ephemeris error and of reference station location error will be similar for the two data types.

The above analysis was performed for the whole view periods of the five GPS satellites in one pass. The ~6 hour data span[2] provides sufficient dynamics needed by the DDGD data. The DDGR data possess intrinsic geometrical information that yields a geometric solution from instantaneous measurements. An analysis was performed for such instantaneous determination of the DSS 14-DSS 13 baseline. The results are shown in Fig. 9. Here the dominating error source is the 10 cm DDGR data noise. This is a result of reducing the number of data points from 93 into 4. A factor of $(4/93)^{1/2}$ reduction in data noise would reduce its effects to a level comparable to the long-arc solutions, shown side-by-side in Fig. 9, indicating the geometrical strength of the DDGR measurements. The effects of other error sources are also comparable with the long-arc solutions. Preliminary test results (Ref. 15) for the same baseline are also shown in Fig. 9 for comparison. The baseline component errors of these short-arc (~1.5 hours) tests lie between the long-arc solutions and the instantaneous solutions.

VI. Summary and Conclusions

Covariance analysis has been performed for the determination of baseline vectors observing the GPS satellites. Doubly differenced data were used to eliminate clock errors. If reasonably stable clocks are used at both ends of the baseline, singly differenced data can be used with periodic adjustments of clock parameters (Ref. 14). Such singly differenced data will possess better information content due to better geometry, because the need for simultaneous observation of at least two

[1] Because of being derived from RF carriers which have much shorter wavelength (19 cm) than the 30-m range-code wavelength.

[2] At least two GPS satellites need to be in view simultaneously to form the doubly differenced data.

GPS satellites is now removed. The frequency at which the clock parameters are adjusted depends on the stability of the clocks.

Even though the analysis has been performed using only the five GPS satellites currently in orbit, the results are believed to be not much inferior to the case using a full constellation of 18 satellites. This full constellation will provide world-wide, continuous 24-hour observations instead of the periodic U.S. coverage with the five satellites.

Precision determination of baselines ~1000 km in length is of particular interest for geodetic applications involving a world-wide station net. A sub-decimeter accuracy over such baselines can be achieved if the GPS ephemerides are known to 1 meter (RSS) and the zenith tropospheric delays to 2 centimeters. Further improvement in baseline vector determination will also require lower data noise and reference station location errors: For centimeter determination of 1000 km baselines, the data noise has to be kept to a 1 cm level, calibration of zenith troposphere to 0.1 cm, reference station location to 5 cm, and GPS ephemerides to 20 cm (rss). The sensitivities to these error sources may be reduced if measurements are made with more than one reference station, thus relaxing the requirements for their precision calibration.

Acknowledgments

The author wishes to thank Claude Hildebrand and Larry Young for their valuable comments. Claude also suggested the use of full covariance matrices to represent the GPS ephemeris errors.

References

1. Fanselow, J. L., et al., "The Goldstone Interferometer for Earth Physics," Technical Report 32-1526, Vol. V, Jet Propulsion Laboratory, Pasadena, Calif., pp. 45-57, Oct. 1971.

2. Thomas, J. B., et al., "A Demonstration of an Independent-Station Radio Interferometry System with 4-cm Precision on a 16-km Baseline," *J. Geophys. Res.*, Vol. 81, No. 5, pp. 995-1005, Feb. 10, 1976.

3. MacDoran, P. F., et al., "Mobile Radio Interferometric Geodetic Systems," Proc. 9th GEOP Conf., *An Int. Symp. on the Applications of Geodesy to Geodynamics*, Oct. 2-5, 1978.

4. Niell, A. E., et al., "Comparison of a Radio Interferometric Differential Baseline Measurement with Conventional Geodesy," *Tectonophysics*, Vol. 52, pp. 49-58, 1979.

5. Sovers, O. J., et al., "Determination of Intercontinental Baselines and Earth Orientation Using VLBI," *TDA Progress Report 42-71*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 1-7, Nov. 15, 1982.

6. Fell, P., "Geodetic Positioning Using a Global Positioning System of Satellites," *IEEE Position Location and Navigation Symp. Record*, pp. 42-47, Dec. 1980.

7. MacDoran, P. F., et al., "SERIES: Satellite Emission Range Inferred Earth Surveying", *3rd Int. Geod. Symp. on Satellite Doppler Positioning*, Las Cruces, N.M., Feb. 1982.

8. Ondrasik, V. J., and Wu, S. C., "A Simple and Economical Tracking System with Subdecimeter Earth Satellite and Ground Receiver Position Determination Capabilities," *3rd Int. Symp. on the Use of Artificial Satellites for Geodesy and Geodynamics*, Ermioni, Greece, Sept. 20-24, 1982.

9. Wu, S. C., and Ondrasik, V. J., "Orbit Determination of Low-Altitude Earth Satellites Using GPS RF Doppler," *IEEE Position Location and Navigation Symp. Record*, pp. 85-91, Dec. 1982.

10. Spilker, J. J., "GPS Signal Structure and Performance Characteristics," *J. Inst. Navigation*, Vol. 25, No. 2, pp. 121-146, 1978.

11. Berman, A. L., "The Prediction of Zenith Range Refraction from Surface Measurements of Meteorological Parameters," Technical Report, 32-1602, Jet Propulsion Laboratory, Pasadena, Calif., July 15, 1976.

12. Thuleen, K. L., and Ondrasik, V. J., "The Repetition of Seasonal Variations in the Tropospheric Zenith Range Effect," Technical Report 32-1526, Vol. VI, Jet Propulsion Laboratory, Pasadena, Calif., pp. 83-98, Dec. 15, 1971.

13. Armstrong, J. W., and Sramek, R. A., "Observations of Tropospheric Phase Scintillations at 5 GHz on Vertical Paths," *Radio Science*, Vol. 17, No. 6, pp. 1579-1586, Nov.-Dec., 1982.

14. Wu, S. C., "Error Estimation for ORION Baseline Vector Determination," *TDA Progress Report 42-57*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 16-31, June 15, 1980.

15. Young, L. E., "SERIES Proof of Concept Test Results," presented at the 5th Annual NASA Geodynamics Program Conf., Washington, D.C., Jan. 24-27, 1983.

Table 1. GPS and baseline geometry

GPS:	Semi-major axis	a = 26,560 km
	Eccentricity	e = 0
	Inclination	I = 63.4°
	Argument of perigee	ω = 0°
	Mean anomaly	M = -66.97° 45.70° -2.40° -90.85° 1.42°
	Longitude of ascending node	Ω = 121.34° 0.05° 121.37° 121.65° 0.31°
	Epoch 1982/05/09 0000 hr UT	
Baselines:	Goldstone local	(150 m north-south)
	DSS 14 – DSS 13	(22 km)
	JPL – Palos Verdes	(58 km)
	DSS 14 – JPL	(181 km)
	JPL – OVRO	(338 km)
	Haystack – OVRO	(4000 km)

Table 2. Error models

Data noise	DDGR 10 cm at 5-min intervals DDGD 0.05 mm/s at 5-min count time
GPS ephemeris	10 m rss over entire view periods
Zenith troposphere	$B < 80$ km, 5 cm at zenith common between stations $0.2\,B^{0.7}$ cm difference between stations $B > 100$ km, 5 cm at each end of the baseline (uncorrelated)
Reference station location	1 m each component

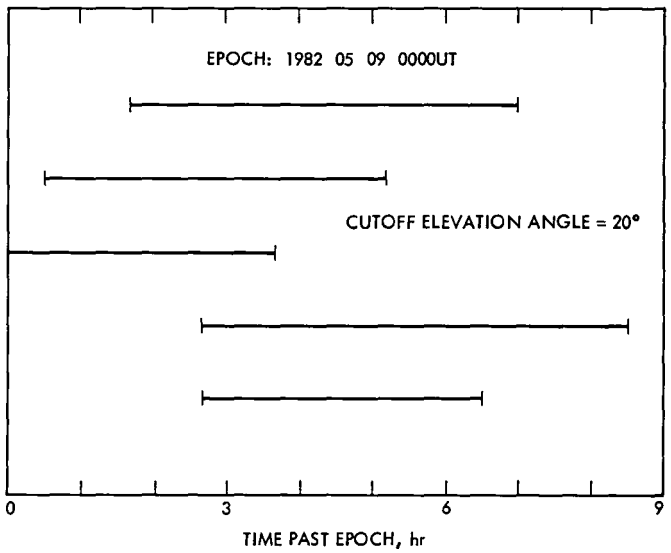


Fig. 1. View periods of five GPS satellites from a station in California

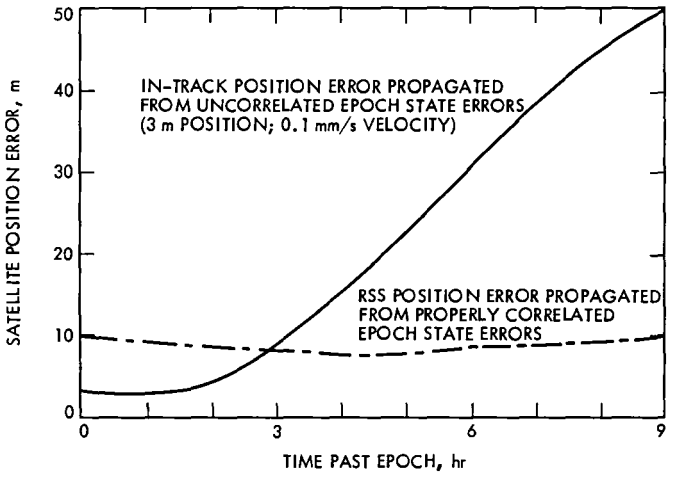


Fig. 2. State error propagation of a GPS satellite

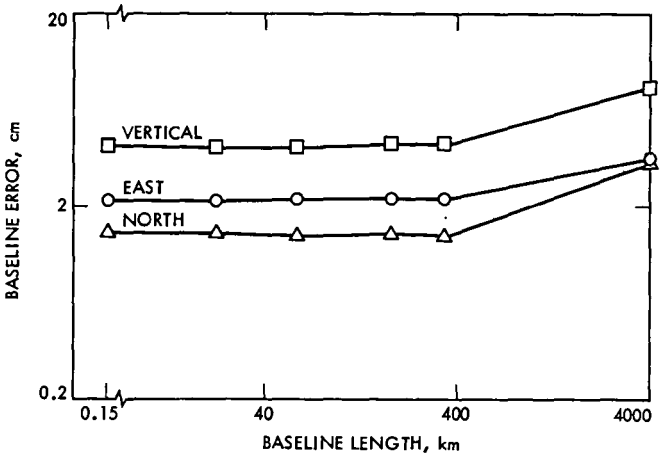


Fig. 3. Effects of DDGR data noise on baseline determination

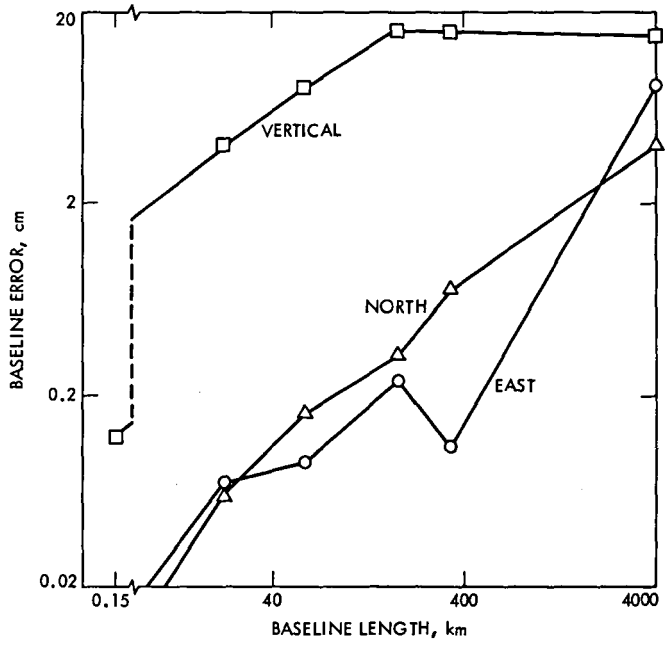


Fig. 4. Effects of tropospheric delay error on baseline determination

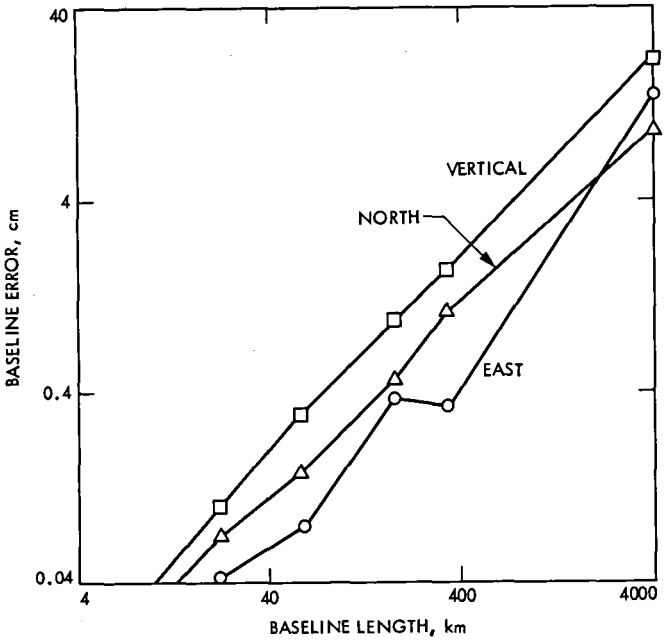


Fig. 5. Effects of reference station location error on baseline determination

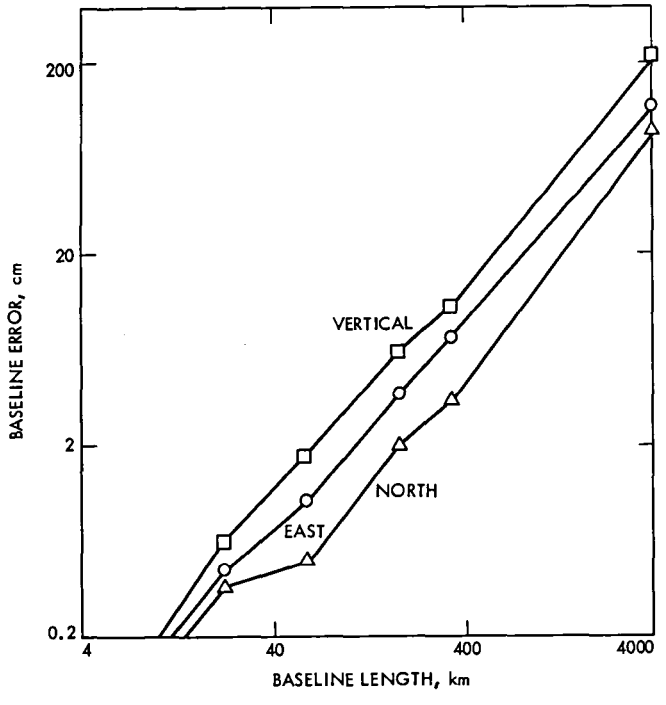


Fig. 6. Effects of GPS ephemeris error on baseline determination

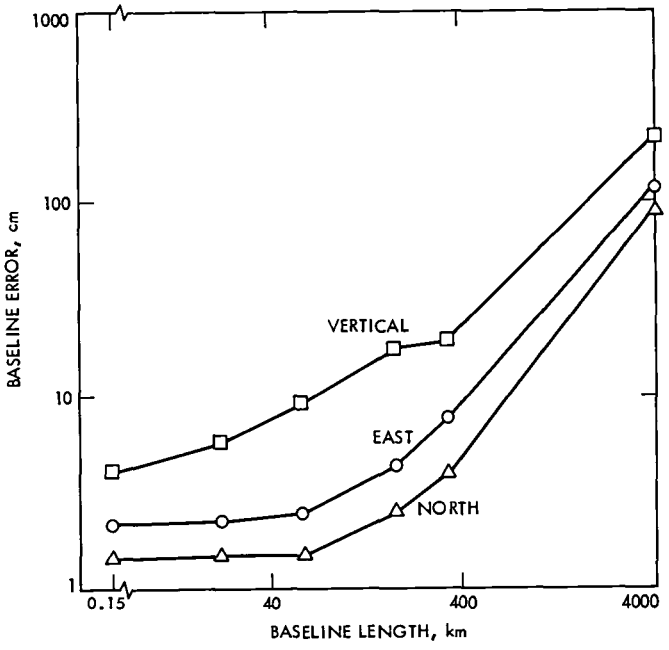


Fig. 7. Total baseline determination error

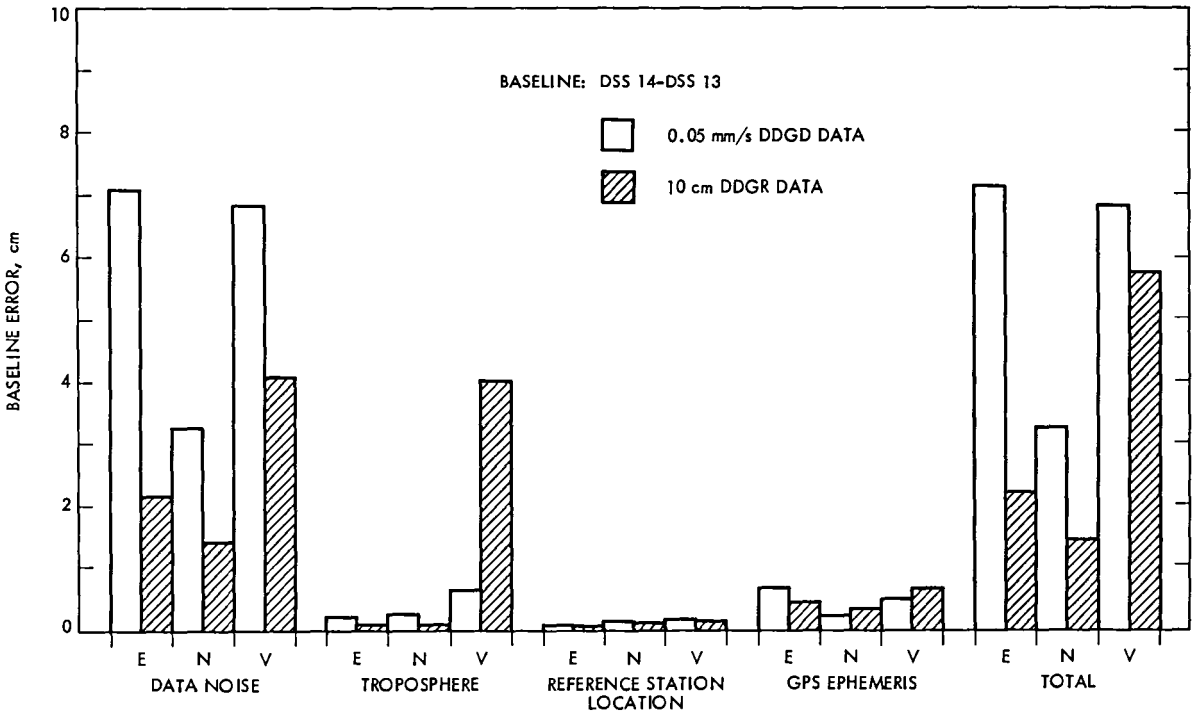


Fig. 8. Comparison of DDGD and DDGR data for baseline determination

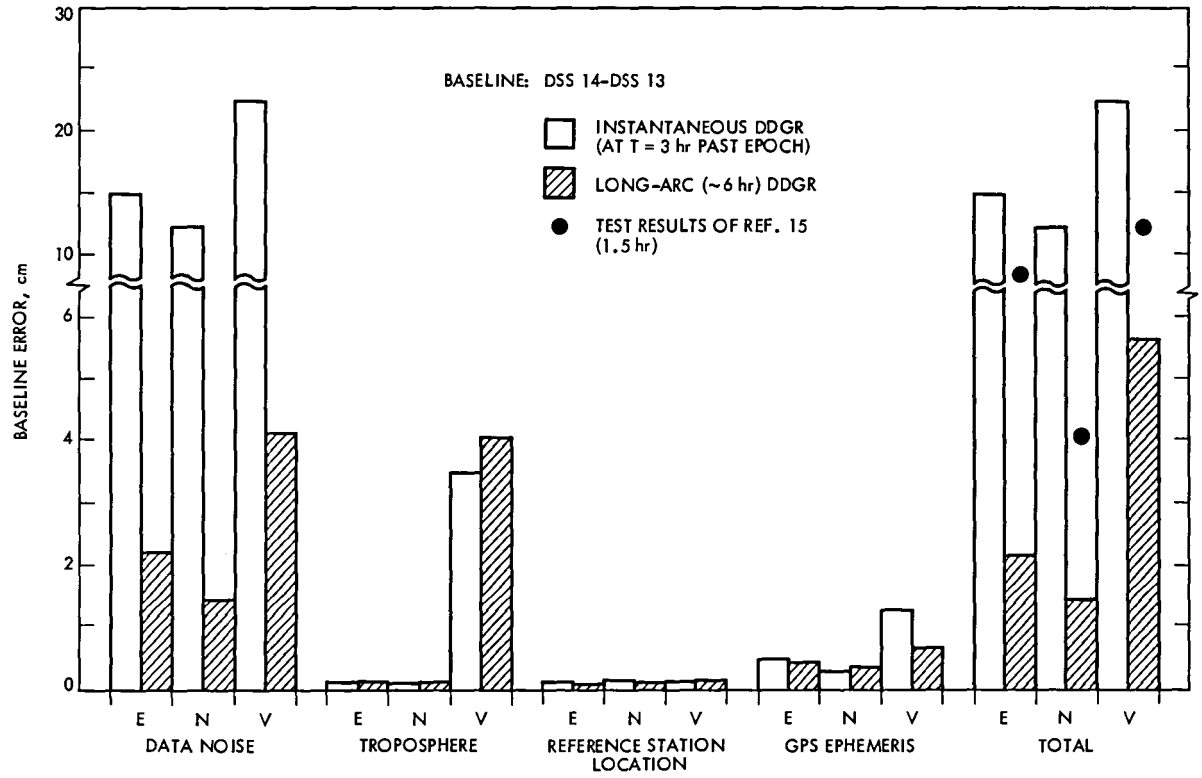


Fig. 9. Comparison of instantaneous DDGR and long-arc DDGR for baseline determination

RF Design and Predicted Performance for a Future 34-Meter Shaped Dual-Reflector Antenna System Using the Common Aperture X-S Feedhorn

W. F. Willams
Radio Frequency and Microwave Subsystems Section

The Networks Consolidation Program (NCP) will utilize 34-meter shaped dual-reflector Cassegrain antennas. This article discusses the shaping calculations, the X/S-band feedhorn to be used, and the predicted RF performance of this antenna system.

I. Introduction

The Networks Consolidation Program (NCP) will include at least two new 34-meter high-aperture-efficiency shaped dual-reflector antennas. This report presents the results of the geometric optics design which determines the dual reflector coordinates of the system and the calculated RF performance values of the antenna when using the recently developed DSN common aperture X-S feedhorn.

The functional requirement for the new NCP antennas is to provide a receive-only capability within modest bandwidths at X-band and S-band. However, there may be future requirements for uplink transmission in one or both bands because of anticipated new programs, e.g., the Galileo and other missions. To cover this eventuality, the microwave subsystem feedcones for the antennas were sized for various future expansions accommodating these potential transmit and other requirements. The microwave hardware, i.e., the feedhorn and S-band combiner, cannot easily be broad-banded and field retrofitted to include these additional functions. Therefore, since the expansions were anticipated, a basic transmit capability and other details have been engineered into the microwave subsystem. The expansion is fairly routine for the X-band portion of the horn alone, but the design for the broad-banding of the S-band injection was particularly vexing.

The objective of the 34-meter design was to maximize the gain/noise temperature (G/T) ratio of the antenna, the so-called figure of merit. A first step in achieving the maximum gain (high aperture efficiency) is to obtain a uniform illumination across the reflector aperture which maximizes illumination efficiency, while at the same time utilizing techniques which maximize all other efficiencies, i.e., reducing, to near elimination, forward spillover past the subreflector and reducing as best as possible the rear spillover (which significantly reduces system noise). This is best achieved (within the present state of the art) by using a specially shaped dual-reflector system. These special shapes are obtained by altering the hyperboloid contour so that the particular feedhorn pattern that is used will be transformed from a usual beam pattern to a basically uniform pattern when scattered from the altered subreflector. This is required for high main reflector illumination efficiency. This process is known to introduce illumination pattern phase departures from the spherical case; main reflector alteration is therefore employed to produce the final required phase uniformity.

The final shaping was determined using the X-band measured horn pattern (8.45 GHz) of the DSN developed X-S common aperture feedhorn. Final predicted values of RF performance were calculated by using the measured pattern data in X- and S-bands, and using theoretical predictions of the scattering from the subreflector. An attempt will be made to estimate certain RF efficiencies which are totally dependent upon mechanical design and construction. These will be the spar (quadripod) blockage efficiency and surface tolerance efficiency. Maximum and minimum values will be suggested for these efficiencies to permit the reader to place bounds on the estimates of final performance.

II. X-S Common Aperture Feedhorn and Combiner

In mid calendar year 1976 (early FY 77) a program was initiated to develop an S- and X-band feedhorn with objectives as follows:

(1) Obtain a centerline symmetric unit to replace present asymmetric simultaneous S/X reflex (dichroic plate) DSN feed systems and thereby

(2) Eliminate the dichroic plate and further optimize X-band performance, with degradation of S-band performance allowed if necessary, and

(3) Plan for the future capability of high power transmission in both X- and S-bands.

Shortly thereafter the present concept was conceived for the dual-band common aperture feedhorn. The major instigation for the concept was obtained from a paper by Jeuken and Vokurka (Ref. 1). In essence, that paper recalls that the corrugation depths for a corrugated horn need to be between $\lambda/4$ (λ = wavelength) and $\lambda/2$ to support the proper HE_{11} waveguide mode. It follows that any such corrugations would be odd multiples of these depths within certain other frequency bands as well. A careful choice of depth is then made to obtain operation within S-band and X-band with depths greater than $\lambda/4$ and less than $\lambda/2$ in S-band and greater than $5\lambda/4$ and less than $3\lambda/2$ in X-band. Thus a sort of "harmonic" operation of the corrugation depth is effected. As such a feedhorn with fixed flare angle is made longer, hence with larger aperture, a point is reached when further increase does not increase horn gain or reduce an associated beamwidth, e.g., 20 dB beamwidth. Details of pattern shape will differ with frequency, but not the gain or this beamwidth. One might call this "saturated gain" operation. Further detail of this approach may be found in Refs. 2 and 3.

As mentioned in the introduction, injecting or extracting S-band from this horn presented a major problem. A solution was found by feeding the horn at a sufficiently large horn diameter (above waveguide cutoff) region from a surrounding radial line which carries the S-band. S-band is injected into this radial line from four orthogonally located peripheral feed points excited in a 90° phase progression to develop circular polarization. The radial line carries two radial rejection chokes which prevents X-band from propagating within the S-band injection device, now termed the X/S-band combiner/separator. The system works very well in X-band since no noise temperature increase was noted when compared to the DSN standard X-band 22-dB horn, which has no S-band operation. The major shortcoming of the first generation combiner was its narrow S-band bandwidth, making it useful in a receive-only system. The problem has been overcome in a second generation combiner. This combiner is discussed in further detail in Ref. 4. A patent was issued for the horn combiner system (Ref. 5).

The final common aperture feedhorn and combiner for NCP is now complete. S-band bandwidth problems and a slight X-band moding problem were solved. Moding was better understood due to work by Thomas (Ref. 6), and the combiner frequency bandwidth was increased by increasing the height of the radial line. A paper (Ref. 7) has been presented which describes the techniques employed to achieve the final performance. A detailed description of the development with photographs and measured radiation patterns is given in Ref. 8. This reference details the predicted performance at DSS 13, using the 26-meter parabolic reflector. The microwave subsystem expansions envisioned for a "4-function feedcone," to cover anticipated future requirements as mentioned above, have already been largely incorporated, on an R&D basis, into the DSS-13 feedcone (Ref. 10). This feedcone is the first simultaneous dual-band receive/dual band transmit DSN system.

III. The Shaped Dual-Reflector System

Shaped dual-reflector antenna systems, sometimes called specially shaped Cassegrain antenna systems, have been in use for approximately 15 years. The first one to be placed in service was designed and built by the Philco-Ford Western Development Laboratory and located in Italy for the Intelsat service. This antenna had an aperture of 30 meters. Since that time a great many more antennas of varying sizes have been built by various companies and placed in service throughout the world.

The shaping consists of slight distortions of the usual hyperboloid subreflector of a Cassegrain system so that the feedhorn pattern can be transformed into a nearly uniform illumination across the main aperture. In so doing the uniform phase pat-

tern, normally present from the hyperboloid, is destroyed. This uniform phase is then recovered by slight (and similar) distortions of the usual paraboloid. The resulting uniform distribution of amplitude and phase gives the maximum possible illumination efficiency available from the given aperture size.

Since the shaping transforms a feedhorn pattern, it follows that a particular feedhorn pattern must be used to obtain the final reflector shape. Geometric optics is used to solve the problem, using only equal path lengths and Snell's law of reflection, so frequency of operation does not enter into the solution. Therefore, any feedhorn pattern at any frequency and reasonably uniform phase will be transformed to an illumination that will be distorted to an extent dependent upon the similarity of its pattern to the pattern that was used in basic shaping design. The geometric optics determined main reflector will then recover the uniform phase required for an antenna aperture. It must be mentioned that the standard paraboloid-hyperboloid Cassegrain system is only a special case of the general shaped dual-reflector antenna. In that case, the transformation is 1 to 1. That is, the feedhorn pattern is unchanged; the illumination becomes that of the feedhorn used in a prime focus system of greater focal length, modified slightly by finite reflector diffraction detail.

The JPL dual-reflector shaping software was prepared in about 1967 and, although exercised often, has never been used for the design of a system that was fabricated. Experimental or test designs of systems have been performed for comparison to the results as computed by two other organizations. The result has indicated that the JPL program compares very well and can be used with confidence.

A new feature has been added to the JPL program. (Ref. 3, p. 10.) This feature eliminates much of the illumination in the central region, hence reducing the subreflector blockage and improving blockage efficiency. Thus, the feature synthesizes what previously was termed a "vertex matching plate," but without an accompanying phase distortion.

As mentioned above, the common aperture feedhorn has been completed and the radiation patterns measured. The measured pattern at 8.450 GHz has been used to calculate the special shapes for the 34-meter NCP antennas. In performing this calculation, the distance from aperture plane to main reflector (quasi-paraboloid) vertex was chosen as a design parameter. This was done to match the 26-meter antennas of the DSN, and so the special shape solutions closely approximate the 26-meter paraboloid contour. The result was a feedhorn focal point location at 193.5 inches (4.915 meters) from the quasi-paraboloid vertex. Hence the new feedcone is compatible with any other DSN antenna for maximum flexibility.

Figure 1 is a control design outline drawing of the shaped antenna configuration. Note that the geometric ray from the quasi-hyperboloid edge does not go through the quasi-paraboloid edge, but instead intercepts at 645 inches (16.383 meters) from the Y-axis. The reason for this choice is as follows: The feedhorn scattered energy from the quasi-hyperboloid does not fall abruptly to zero at the angle θ_1 (Fig. 1), but instead tapers rapidly to a low level. The angle θ_1, and hence the 16.383-meter dimension, is chosen so that the intensity at θ_2 may be at a very low level relative to the central region of the main reflector, and the resulting rear spillover noise contribution becomes acceptably small. This results in a slightly lower illumination efficiency and hence antenna gain, but the significant reduction in noise from rear spillover allows an optimum G/T ratio.

Also note from Fig. 1 that an equivalent F/D is listed. It must be emphasized that this antenna has no unique focal length "F," but instead only a variable focus position which is a function of any particular ray from feed to subreflector to main reflector. The equivalent F/D is listed to serve as a description of the main reflector depth or the dimension from aperture to vertex for purposes of quick comparison to standard Cassegrain systems.

Complete Specification Control drawings for the dual-reflector system including coordinates of both reflectors are presented in JPL drawings 10097335 and 10097336.

As mentioned above, the shaping solution was performed to obtain a result that was a "best fit" to existing 26-meter paraboloid antennas. This was done so that individual 26-meter panels might be used to approximate the required shaped surface, hence potentially saving costs in the shaped surface fabrication. However, it was later decided to fabricate a complete new antenna structure with all new panels, but the original calculated contours have been retained.

IV. Predicted RF Feed System Performance of the New Antennas

The predictions of RF performance for the new NCP antennas were based upon the theoretical scattering (computer programs) of the measured common aperture X-S horn patterns from the shaped subreflector surface. Results will be presented for frequencies in both X- and S-band.

A JPL computer program has been prepared which calculates the efficiencies of a paraboloidal reflector system based upon its illumination pattern. All the calculated efficiencies remain the same for our quasi-paraboloid relative to a paraboloid system except for phase efficiency. As was mentioned earlier, the scatter pattern from the shaped subreflector has a

nonuniform phase function; the funtion of the quasi-paraboloid is to correct for this to obtain the necessary uniform phase front. This is the only function or purpose of the modification to the paraboloid. Therefore, another computer program has been prepared which performs this phase adjustment just as the quasi-paraboloid does. The illumination pattern for efficiency calculation then has the near uniform phase function, but normal feed phase variations and hence diffraction phase characteristics have been maintained. This phase corrected illumination pattern is then used in the efficiency program.

Figures 2, 3, 4 and 5 present subreflector scattered patterns of the measured feedhorn patterns at the four noted frequencies. Table 1 lists the resulting calculated efficiency elements comprising the RF feed system performance. An additional efficiency of 0.97 has been included to represent an estimate of dissipative loss and loss due to voltage standing wave ratio. It must be emphasized that no allowance has been made for the structural/mechanical losses mentioned above.

A noise level in kelvins (K) is listed beside values of rear spillover. This is an estimate of noise contribution from the spillover energy that will come from the earth source. The estimate is made as follows: (1.0 − η (rear spillover)) is the fractional noise energy from the earth that enters the antenna feed system and receiver. The estimate made is that the average blackbody radiation is such as to make the temperature 240 K, instead of the 300 K maximum. Therefore, 240 K is multiplied times this fraction to obtain values of temperature directly attributable to the "hot" earth.

It now becomes possible to maximize G/T for different aperture illumination angles, again considering only the RF feed system. These RF values will be reduced for a final estimate by the values of gain reduction attributable to surface tolerance and spar blockage. For the G/T estimate it is assumed that the noise from rear spillover is reduced by 50% at a 30-deg elevation angle since much less earth is then seen by the rear spillover. Table 2 presents the results of these calculations. This table shows the result of making the quasi-hyperboloid edge ray incident upon a point well within the quasi-paraboloid aperture edge, i.e., $\theta_2 > \theta_1$ (see Fig. 1). If the design were made so that $\theta_1 = \theta_2$ (ray from edge to edge) the rear spillover noise would be from 4 K to 8 K, a value large enough to significantly reduce G/T from the potential maximum. Figure 6 presents the information in graph form, indicating a region of optimum operation.

Final secondary patterns have also been calculated. This is done by using the phase altered scatter patterns of Figs. 2 and 4 as illumination functions for a true paraboloid. The paraboloid is chosen to have the same total illumination angle as the shaped system, i.e., 2 X 72.79 = 145.58 deg. These calculated patterns are depicted in Figs. 7 and 8. It must be pointed out that these are "best possible" patterns since neither surface tolerance nor spar blockage has been considered, and each will increase side lobe level and decrease gain somewhat. Beamwidth prediction should be quite close.

V. Final System Performance Based Upon Blockage and Surface Tolerance Assumptions

A short discussion of surface tolerance efficiency by Ruze and of JPL experience with spar blockage efficiency is given in Ref. 9. It will be the purpose of this final section to use extremes of both efficiencies and calculate the range of final gain levels that can be expected. It is not a purpose of this report to assign or predict these mechanical results, but only to direct attention to their final effect.

The rms surface error ϵ is expected to be as good as 0.5mm and no worse than 1.0mm. Using the Ruze formula, the resulting efficiencies may be as follows:

	Surface efficiency	
ϵ (mm), rms	8.45 GHz	2.295 GHz
1.0	0.8822	0.9908
0.5	0.9692	0.9977

Area blockage due to the feed support spars (quadripod) is expected to be between 6% and 8%. From JPL experience, the final blockage efficiency will be as follows, using

$$\eta \text{ (spar blockage)} = (1 - 1.2(Ap))^2$$

with Ap representing the geometric shadowing percentage by the spars.

Fractional area blocked	η (spar blockage)
0.06	0.861
0.08	0.817

When applying these extremes to the previously determined values of RF performance, the final gain results will be as tabulated in Table 3.

A most significant result to note is that X-band performance may improve almost a full decibel by careful engineering design of a best surface tolerance and minimum quadripod blockage. The design is seen to yield 68% X-band aperture efficiency within the very low noise design constraint under best conditions while providing simultanous 66% S-band efficiency.

References

1. Jeuken, E. J., and Vokurka, V. J., "Multi-Frequency Band Corrugated Conical Horn Antenna," *1973 European Microwave Conference Proceedings*, Vol. 2, Brussels University, Brussels, Belgium, Sept. 4-7, 1973.

2. Williams, W. F., "A Prototype DSN X-S Band Feed: DSS 13 First Application Status," *DSN Progress Report 42-44*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 98-103, Apr. 15, 1978.

3. Williams, W. F., "DSN 100-Meter X- and S-Band Microwave Antenna Design and Performance," Publication, 78-65, Jet Propulsion Laboratory, Pasadena, Calif., pp. 13-14, Aug. 1978.

4. Williams, W. F., "A Prototype DSN X- and S-Band Feed: DSS 13 Application Status (Second Report)," *DSN Progress Report 42-47*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 39-50, Oct. 15, 1978.

5. United States Patent No. 4199764, "Dual Band Combiner for Horn Antenna," Apr. 22, 1980, W. F. Williams and Seymour B. Cohn.

6. Thomas, B. Mac A., and Minnett, H. C., "Propagation in Cylindrical Waveguides with Anisotropic Walls," Publication RPP 1346, Division of Radiophysics, CS1R0, Sydney, Australia, Jan., 1977, also in *IEEE Transactions*, AP, Mar. 1978.

7. Williams, W. F., and Withington, J. R., "A Common Aperture S- and X-band Feed for the Deep Space Network," presented at the 1979 Antenna Applications Symposium, U. of Ill., Allerton Park, Sept. 1979.

8. Williams, W. F., and Reilly, H., "A Prototype DSN X/S Band Feed: DSS 13 Application Status (Fourth Report," *TDA Progress Report 42-60*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 77-78, Dec. 15, 1980.

9. Williams, W. F., "LAAS Studies: 26-, 34-, and 40-Meter Elements," *DSN Progress Report 42-51*, Jet Propulsion Laboratory, Pasadena, Calif., p. 158, June 15, 1979.

10. Withington, J. R., and Williams, W. F., "A Common Aperture X- and S-band Four Function Feedhorn," presented at the 1981 Antenna Applications Symposium, University of Illinois, Allerton, Park, Sept. 1980.

Table 1. Final RF efficiency

	8.450 GHz	7.150 GHz	2.295 GHz	2.110 GHz
Rear spillover efficiency η_{rs}	0.9982 (0.4 K)	0.9970	0.9746 (6.1 K)	0.9695
Fwd. spillover efficiency η_{fs}	0.9839	0.9744	0.9100	0.9034
Illumination efficiency η_i	0.9823	0.9781	0.9874	0.9849
Cross-Pol. efficiency η_x	0.9988	0.9990	0.9995	0.9994
Phase efficiency η_{ph}	0.9744	0.9693	0.9383[a]	0.9593[a]
Cent. blockage efficiency η_{cb}	0.9826	0.9815	0.9765	0.9767
Dissipation, VSWR efficiency η_d	0.9700	0.9700	0.9700	0.9700
Total RF efficiency η	0.8949	0.8760	0.7779	0.7835

NOTE: Notice that no surface tolerance efficiency or quadripod blockage efficiency has yet been assigned.

[a]S-band phase center does not coincide with 8.450-GHz phase center, resulting in somewhat poorer phase efficiency at S-band. (Horn location and shaping are based upon the 8.450 GHz phase center.) 2.110-GHz phase efficiency calculates greater than 2.295 GHz because of slight differences in the scattered diffraction pattern.

Table 2. G/T calculations at 8.45 GHz

θ_2 illumination angle, deg	Total RF efficiency	Noise, rear, K (a) Zenith (b) 30° elevation	*Final $T(op)$, K (a) Zenith (b) 30% elevation	G/T, dB	Notes
70.78°	0.8872	(a) 7.54 (b) 3.77	25.04 26.37	55.06 54.84	$\theta_1 = \theta_2$
71°	0.8936	(a) 7.11 (b) 3.55	24.61 26.15	55.17 54.90	
71.949°	0.9052	(a) 1.8 (b) 0.9	19.30 23.50	56.28 55.42	θ_2 for maximum gain
72°	0.9052	(a) 1.6 (b) 0.8	19.10 23.40	56.32 55.44	
72.79°	0.8949	(a) 0.43 (b) 0.22	17.93 22.82	56.55 55.50	Maximum G/T for this antenna
73°	0.8908	(a) 0.25 (b) 0.12	17.75 22.72	56.57 55.50	
74°	0.8648	(a) 0.083 (b) 0.042	17.58 22.64	56.49 55.39	

[a]Total operating noise temp., neglecting rear spillover, is taken as 17.5 K for zenith and 22.6 K at 30° elevation.

Table 3. Final performance estimation

	8.450 GHz efficiencies		2.295 GHz efficiencies	
	max	min	max	min
RF efficiency	0.8949		0.7779	
Spar blockage efficiency	0.861	0.817	0.861	0.817
Surface efficiency	0.9692	0.8822	0.9977	0.9908
Total efficiency	0.7468	0.6450	0.6682	0.6297
Resulting gain limits, dB	68.30	67.66	56.49	56.24

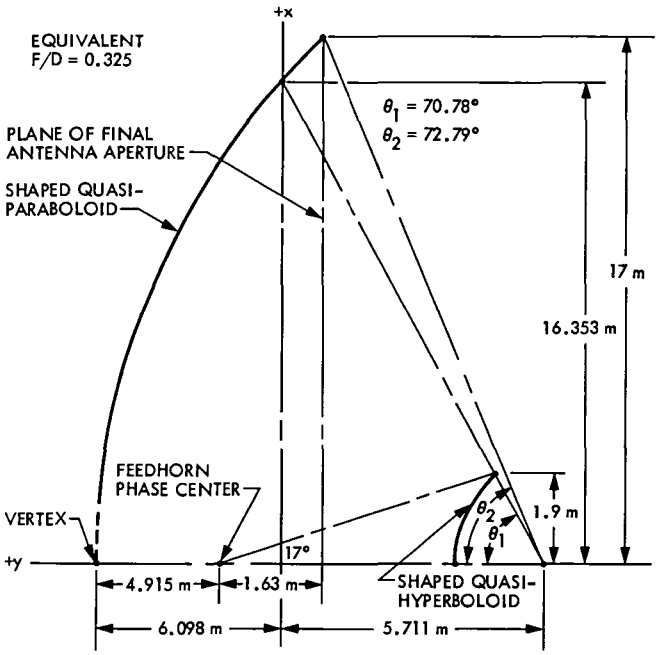


Fig. 1. Design control specification

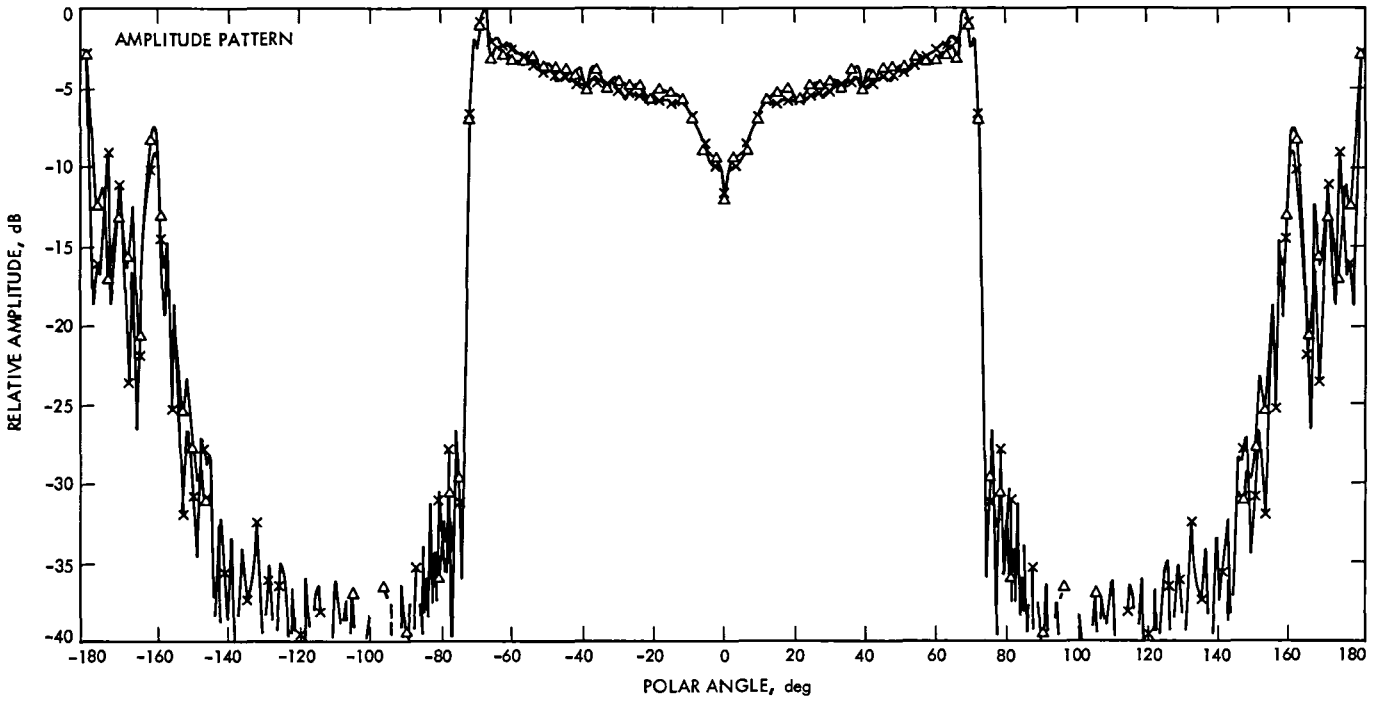


Fig. 2. Shaped subreflector scattering. Mod 2 X/S horn at 8450 MHz scattered from a 34-meter shaped subreflector

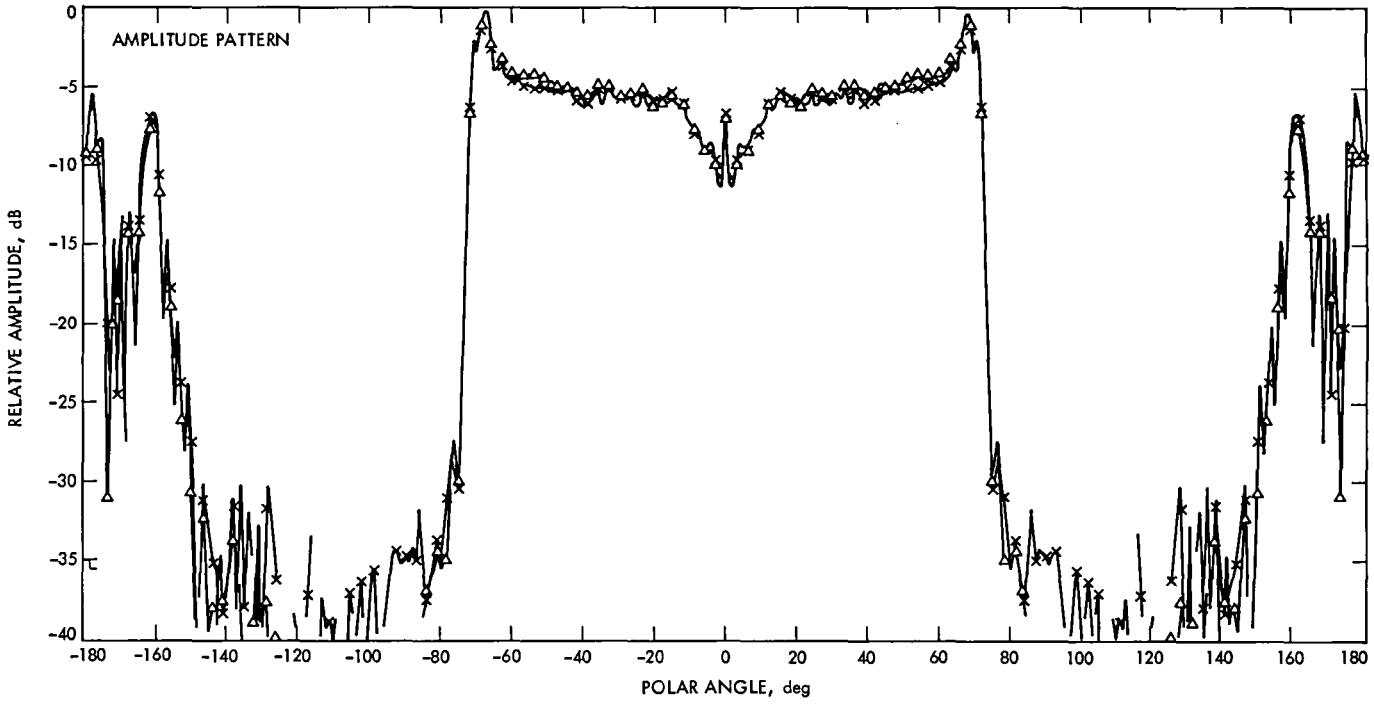


Fig. 3. Shaped subreflector scattering. Mod 2 X/S horn at 7150 MHz scattered from a 34-meter shaped subreflector

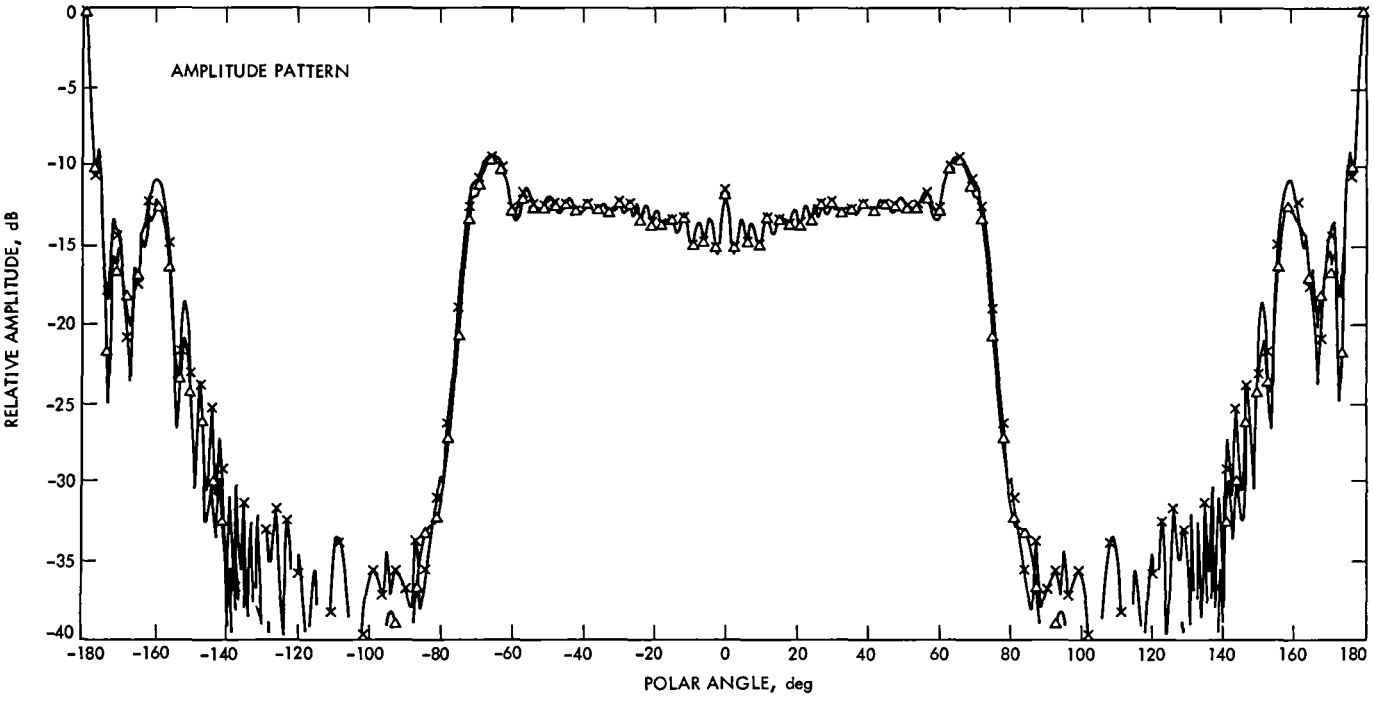


Fig. 4. A subreflector scattered pattern. Mod 2 X/S horn at 2295 MHz scattered from a 34-meter shaped subreflector

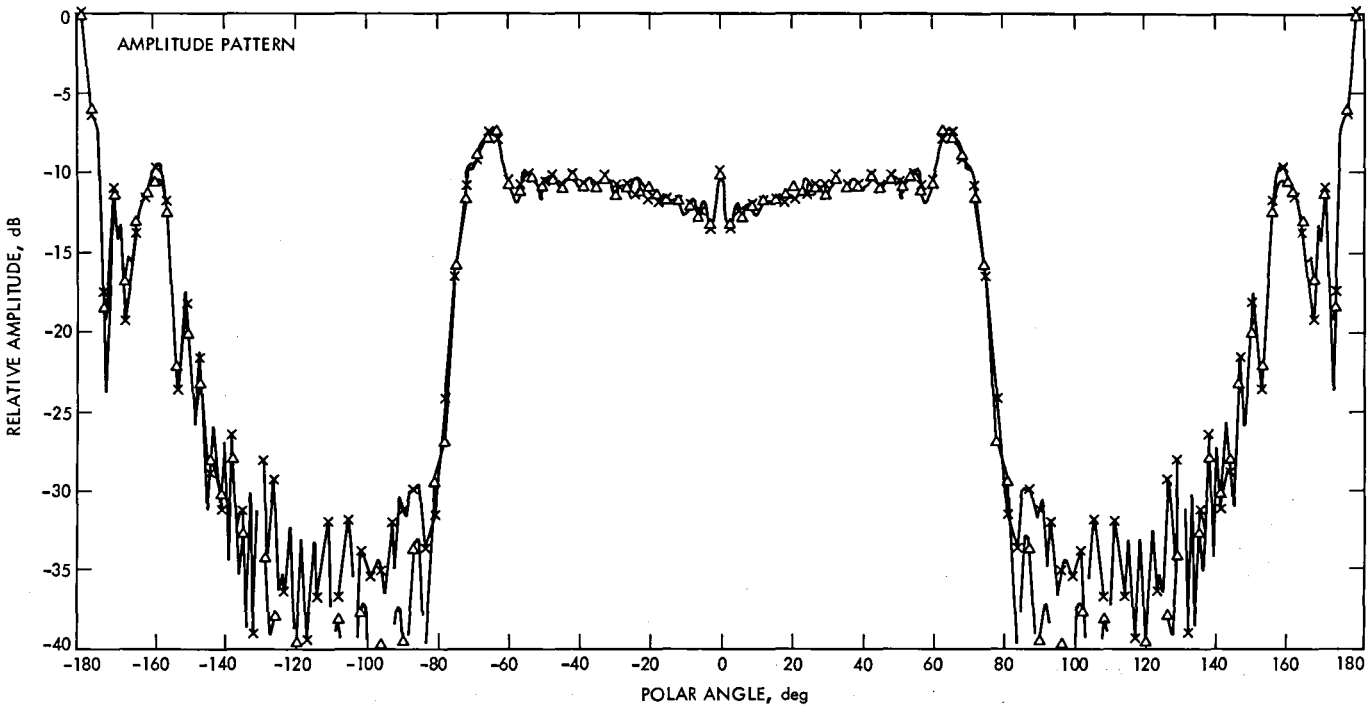


Fig. 5. Subreflector scattering. Mod 2 X/S horn at 2110 MHz scattered from a 34-meter shaped subreflector

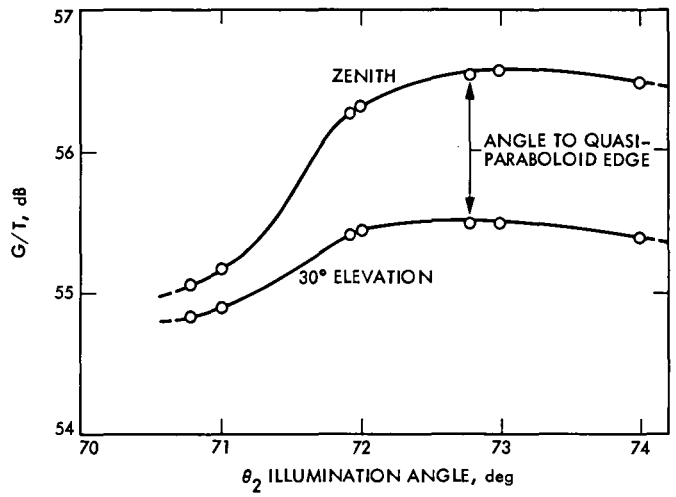


Fig. 6. G/T of system vs edge illumination angle

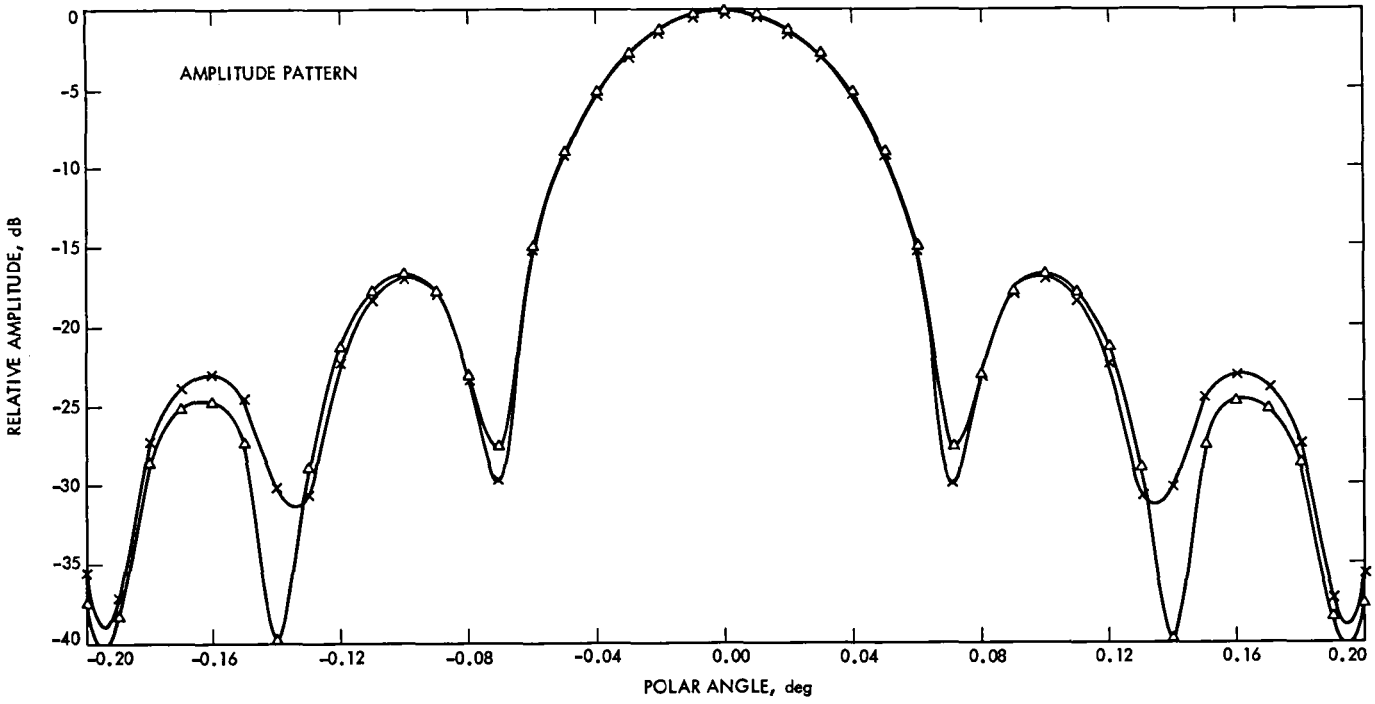


Fig. 7. Secondary pattern scattering. The 8450-MHz real cassegrain pattern after phase correction. Scattered from the equivalent 34-meter paraboloid. F = 454 inches, giving the proper illumination

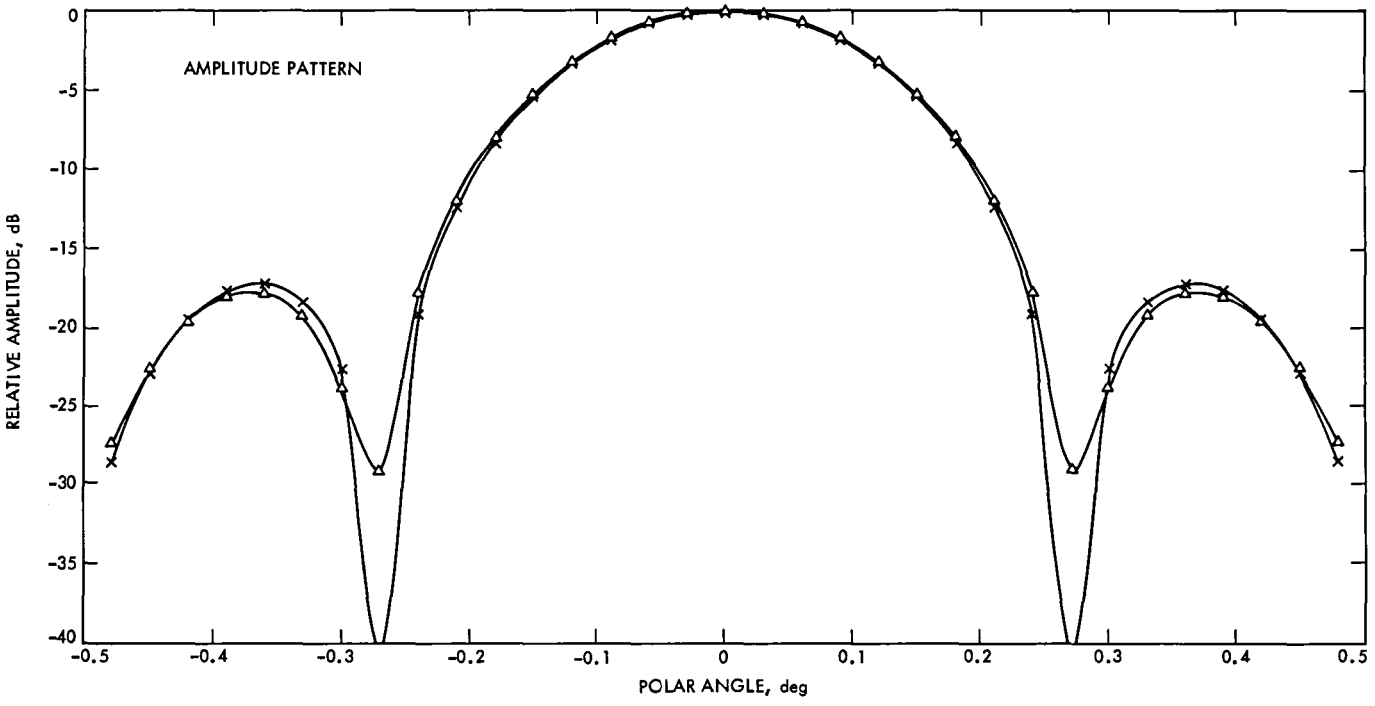


Fig. 8. Secondary pattern scattering. The 2295-MHz real cassegrain pattern after phase correction. Scattered from the equivalent 34-meter paraboloid. F = 454 inches, giving the proper illumination

The Network Information Management System (NIMS) in the Deep Space Network

K. J. Wales
Operations Sustaining Engineering

In an effort to better manage enormous amounts of administrative, engineering, and management data that is distributed worldwide, a study was conducted which identified the need for a network support system. The Network Information Management System (NIMS) will provide the Deep Space Network with the tools to provide an easily accessible source of valid information to support management activities and provide a more cost-effective method of acquiring, maintaining, and retrieving data.

I. Introduction

The NIMS (Network Information Management System) will reside on a Tandem computer system and is being designed to provide distributed administrative data processing. The NIMS will perform the following major functions in the Deep Space Network (DSN): data acquisition, data management, and data distribution.

For several years the DSN has been using computers in a variety of ways to assist in the administrative, engineering and management tasks of keeping the DSN functioning successfully. These various computer applications have evolved sporadically, individually justified on their own merits and with varying degrees of success. They have been implemented on a variety of computers, ranging from the Institutional Computer at JPL through various minicomputers to some recent microprocessor-based systems such as the Automated Office Data Center (AODC).

Although most of the individual efforts have been successful, lack of coordination has contributed to the following state of affairs:

(1) Duplication and inconsistency of data among files.

(2) Inaccessibility of files to those not expert in data processing languages.

(3) Obsolete or inaccurate data.

(4) Inability to move data from one facility to another.

(5) Denial of several applications because they were not individually cost justified.

The Configuration Control and Audit (CCA) project, established in the early 1970s, provided much documentation and groundwork relating to the NIMS project. The CCA project identified the needs for a network support subsystem,

provided the basic tradeoff studies, and conducted a demonstration which showed the validity of the concept (Ref. 1).

A recent study, under the NIMS project, was conducted which consisted of a users survey and an analysis of current needs of the DSN. The study primarily focused on Goldstone Operations (378), Operations Sustaining Engineering (377), and Control Center Operations (371), with a cursory study of Engineering Sections 331, 333, 338, 355, and 430. The Deep Space Stations in Spain and Australia were presented a model of the Goldstone survey for any relevant updates and comments.

The objectives of the NIMS efforts evolved in response to the needs identified from the above study and are outlined below:

(1) Provide an easily accessible source of valid information to support DSN management activities.

(2) Provide a more cost-effective method of acquiring, maintaining, and retrieving data.

II. NIMS Key Characteristics

The NIMS is being designed as a general-purpose computer utility that can be adapted to future needs by specific user application programs. Each of these application programs can exist independently or in concert with one another on the system. Computer tools and services will be provided by vendor-supplied software to support applications programmers, data base administrators, and casual users.

Based primarily on user analysis, the NIMS will feature the following characteristics:

Distributed Processing System. Computational resources are distributed across the network thus reducing communications costs and providing local control and responsibility for the data. This eliminates complete dependence upon a central computer, which can be catastrophic when communication links are down.

Networking Software. Networking features such as remote logon, file access, use of peripherals and services, and message routing through the network are essential to the NIMS distributed environment. Packet switching technology will be used to accomplish the task.

Modularity. The ability to add applications, nodes, terminals, and communications facilities over time are vital to the NIMS success. The ability to connect to existing DSN computer hardware for message and data transfer will be supported.

Utilization of Existing DSN Facilities, Computers, and Services. Consideration is given to existing DSN facilities in supporting NIMS implementation. Existing DSN communication facilities will be used to interconnect NIMS modes. Protocol to interface DSN mini- and microcomputers as well as Telemail and TWX services will be provided.

Database Management System (DBMS). The ability to logically address data, define data dictionaries, and describe "user views" are important capabilities for the system to possess. Because the DSN has many complex data relationships, a network model DBMS is the most satisfactory to implement.

File Management. Key, random, and sequential access methods will be supported.

Programming Tools. BASIC, FORTRAN, and COBOL will be provided for application program development. These languages should interface to the NIMS DBMS and file management system where data generated by them should be compatible between them. A system programming language should be available as well as tools for defining screen formats and menus.

System Performance. Operator response will be between 2-5 seconds, with no response delay greater than 20 seconds. Benchmark analysis will be performed during the vendor evaluation process.

Transaction Processing. The NIMS will support both batch and interactive processing.

Security. The NIMS operating software will provide entry and password access to files and databases to guard against accidental or malicious access to data and application programs. Network security provisions will also be supported.

On-Line Support Facilities. The NIMS will provide facilities that aid in the operation of the NIMS itself, management and maintenance of the database, and implementation of new features. Means are provided for operator visibility into the communications and system status. Diagnostic programs will be available to isolate failing elements.

Availability. As a distributed system, failures will tend to be localized so remaining nodes still function unless requiring data from the failed node. The Tandem Non-Stop II design maximizes individual node availability. The NIMS networking applications depend upon the availability of the NASA Com-

munications (NASCOM) circuit to establish the communications link. These factors make the total availability parameter more difficult to define.

Reliability. Mean Time Between Failure (MTBF) and Mean Time to Repair (MTTR) parameters were criteria applied in the vendor selection process. A MTBF of 750 hours for NIMS node operations is needed for operational satisfaction. The MTTR for a node must be less than 4 hours.

System Integrity. The NIMS, its DBMS and support utilities will provide the means to back up the database and protect it in the event of catastrophic failures and restore it to a previous known state.

III. NIMS Configuration

Several different configurations were studied for the NIMS. Because the NIMS is intended to be a computer-based information management system, the studies were carried out in the two general areas of hardware and database management software. The resulting configuration is discussed below.

A. Hardware

The current administrative data communications arrangements between the complexes depend on a single NASA Communications (NASCOM) circuit. The NIMS application study has shown the need for many diverse activities to be supported. Time-sharing the single line will not provide the required support capability of NIMS nor is it the most cost-effective method of accomplishing NIMS tasks. Accordingly the NIMS has been designed as a distributed data processing system, thereby alleviating the communications circuit time-share problem.

The NIMS is comprised of four nodes. These four nodes are located at the three Deep Space Communication Complexes (DSCCs) and Pasadena. Each of these nodes has one or more terminals connected to it. The Pasadena node will serve as the communications node interconnecting the three regular nodes.

B. Software

The heart of the NIMS system is the software that performs the function of database management. These software packages are commonly referred to as Data Base Management Systems (DBMS). A network model DBMS which directly shows actual relationships and correlations that exist between the data would be the best model to use for DSN data. A relational model DBMS would be the next best choice. Tandem supplies a hybrid relational DBMS that incorporates some of the features of a network model.

This DBMS will perform data and message definitions, terminal and transaction control, data base consistency, enquiry and reporting, and data base maintenance.

Other vendor-supplied off-the-shelf software will provide utilities to monitor the network environment, manipulate files and peripherals, create text and source code, monitor and report individual node statistics, and perform system diagnostic checks.

C. Vendor Selection

Data General, Digital Equipment Corporation, Hewlett-Packard, Prime and Tandem Computer were asked to participate in a competitive evaluation conducted by JPL. Evaluation was conducted in the areas of hardware, software, communications characteristics, reliability, maintainability and contractor support. The Tandem Nonstop II architecture was chosen as best suited to the NIMS project requirements.

IV. First NIMS Application

The Engineering Change Management (ECM) data base is an administrative tool to support management of the engineering change process in the DSN. An upgrade is currently in development (Phase II) to eliminate perceived major shortcomings of the existing ECM Phase I system. Problems of inflexible report writing programs, lack of visibility into details of assessment phase, awkward and inflexible file structure and inability to communicate with other data bases, thus requiring duplicate entries when updating milestones, are to be solved in ECM Phase II. This upgrade will be the first application implemented on the NIMS.

The ECM environment is a distributed one in which portions of the ECM database are placed at the nodes that are using the data. The ECM database is the sum of the individual pieces in the NIMS nodes supporting that application.

Phase II ECM implementation calls initially for a single NIMS node to be installed in the ECM office. Approximately six months after hardware installation ECM Phase II will provide the following services to ECM users: flexible, user-defined reporting system, independence from the Univac 1100/81F, paperless operation, visibility of the assessment process, menu-driven interface for updates and user-friendly operation. ECM Phase II will be implemented in phases as the NIM nodes are installed at the complexes.

V. Summary

The NIMS is designed as a modular computer utility providing tools and services which go beyond mere use of the system. Performance is met by the implementation of application programs to serve user needs. The application programs can exist independently of one another on the system or in concert with one another. Applications programmers, data base administrators, engineers, DSN management and casual users will have at their disposal the computer tools to accomplish whatever job task necessary.

Several applications have been identified for possible future implementation on the NIMS. Databases that can provide benefit to the DSN include financial management, equipment and materials management, maintenance and production control, tracking scheduling and anomaly reporting services.

Reference

1. Bryan, A. I., "A Distributed Data Base Management Capability for the Deep Space Network" *DSN Progress Report 42-33*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 32-36, June 15, 1976.

Evaluation of Antenna Foundation Elastic Modulus

H. McGinness and G. Anderson
Ground Antennas and Facilities Engineering Section

An experiment to measure the elastic deflection of the DSS 14 concrete pedestal under the weight of the antenna was conducted in February 1983 and is compared to a similiar experiment made in 1968. Comparison of the results confirms the decrease in elastic modulus measured on core samples recently taken from the pedestal.

I. Introduction

The 64-m antennas at DSS 14, 43, and 63 have hydrostatic azimuth thrust bearings. In order for the thin oil films, constituting these hydrostatic bearings, to be maintained at safe finite thicknesses, the bearing runner foundation must have a certain minimum stiffness and dimensional stability. Since the foundation pedestal of this antenna design is made of steel-reinforced concrete which is stressed only moderately by the antenna loads, it was supposed that its stiffness and stability would last indefinitely. Recent investigations have disclosed that the DSS 14 foundation is undergoing alkali-aggregate reaction, a phenomenon caused by acidity of the aggregate. This reaction causes the aggregate-cement bond to fail, thus causing a reduction in the strength and in the elastic modulus. Of the three 64-m antennas, only the one at DSS 14 is known to contain aggregate which permits the occurrence of this phenomenon.

II. Description of the Experiments

In 1968 a sensitive and accurate level (trade name Talyvel) was placed on a runner outer alignment lug near trough azimuth 283 deg. (See Fig. 1 for definitions of trough and antenna azimuth angles.) The level was mounted so as to read slopes in the tangential direction. When the antenna azimuth was 157 deg, the left front bearing pad was centered opposite the Talyvel level at trough azimuth 283 deg. At this point the level was set to read zero. Then the antenna was moved counterclockwise 120 deg, with the level readings being made every 5 deg. The antenna was then moved clockwise 120 deg, with the level readings being repeated every 5 deg. This was equivalent to measuring the slopes at various distances from the loaded pad.

The relative vertical deflection between two points can be obtained by numerically integrating the slopes between the same two points. The result was that the relative deflection between the pad center and a point halfway to the center of the next pad, 60 deg away, was 1.07 mm. If the assumption of zero absolute deflection at the midpoint between pads is valid, then the relative deflection between the pad and midpoint is also the absolute vertical deflection at the pad center.

On February 28, 1983, similar experiments were made. Talyvel levels had been mounted on the runner outer alignment lugs near trough azimuth stations 79, 199, 319, and 135 deg. Additionally, levels had been mounted near the outer edges of the top of the concrete pedestal at these same

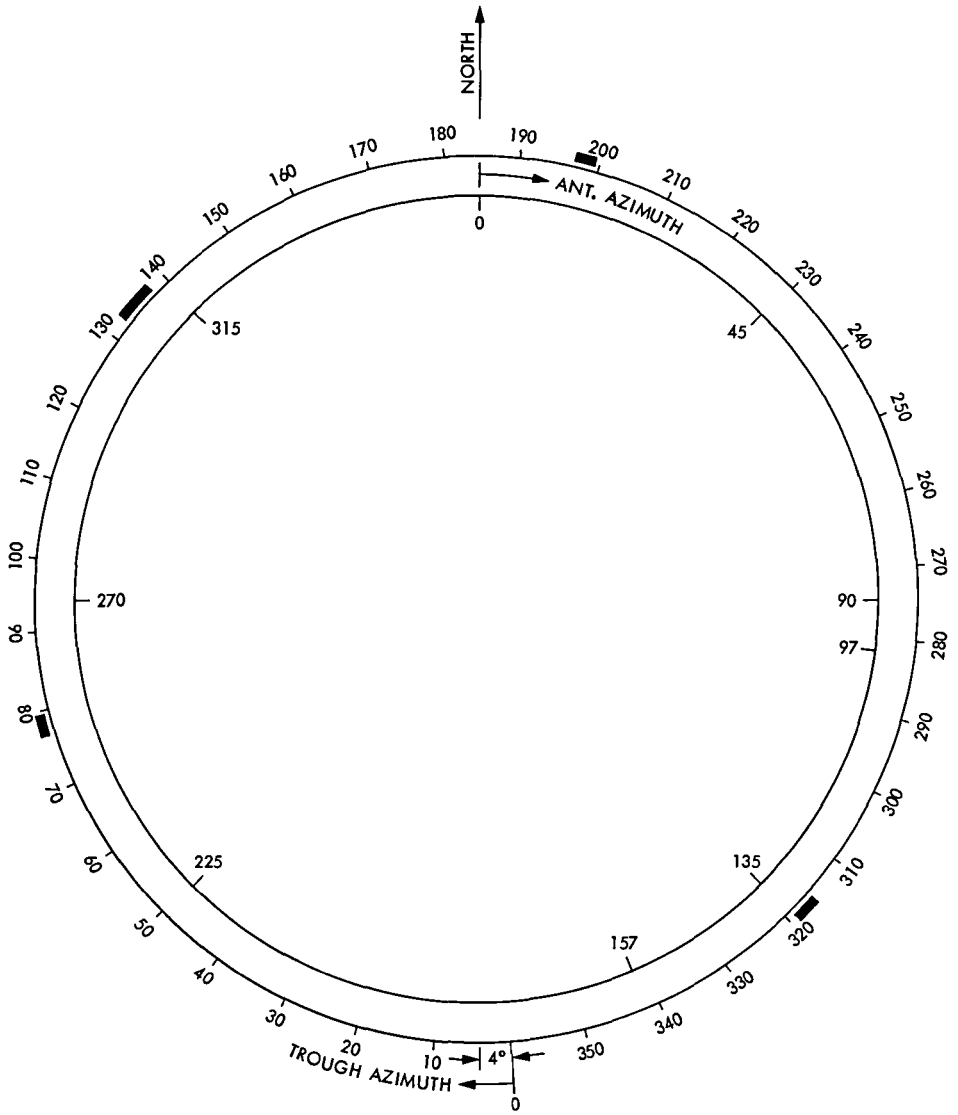
trough azimuth angles. Since the first three of these angles are spaced 120 deg apart, which is also the spacing between the three hydrostatic bearing pads, three sets of data were collected simultaneously. The fourth level position at 135 deg trough azimuth was so placed because of the history of much runner shimming at this station. Unfortunately the level on the alignment lug at station 319 deg did not work properly, but all the other levels performed satisfactorily and the repetition errors were small. Because of symmetry it is believed that the deflections derived from these levels can legitimately be compared to the 1968 deflection. The three relative deflections derived are 1.59, 1.59 and 1.80 mm, the average value of which is 1.66 mm. When the pad centerline was within 10 deg of the level, readings were taken in 1-deg increments. When the pad was more than 10 deg from the level, readings were taken in 5-deg increments.

It is believed that the errors on these deflections do not exceed 10%.

III. Conclusion

The relative deflection between points of an isotropically elastic structure is inversely portional to the elastic modulus. Since the present average relative deflection of 1.66 mm is significantly greater than the 1968 relative deflection of 1.07 mm, the present elastic modulus is 1.07/1.66 or 64% of the 1968 elastic modulus.

Several core samples have been taken from the pedestal within the last year. The measured elastic moduli of eight samples taken in December 1982 had an average value of 2.3×10^{10} pascals and a standard deviation of 0.5×10^{10} pascals. Soon after construction of the pedestal in 1964 the elastic modulus was 3.8×10^{10} pascals. It is believed that there had been no deterioration of the modulus by 1968. The ratio of the current modulus to the 1964 modulus is 2.3/3.8 or 60%. This core sample modulus ratio of 60% checks fairly well with the measured deflection ratio of 64%.



$A_{TROUGH} = A_{ANT} + 360\ n - 174°$, WHERE n = 0 OR 1

Fig. 1. Location of level instruments on runner outer lugs

Techniques for Analysis of DSN 64-Meter Antenna Azimuth Bearing Film Height Records

R. Stevens, Chief Engineer
Telecommunications and Data Acquisition

C. T. Quach
Telecommunications Science and Engineering Division

The DSN 64-m antennas use oil pad azimuth thrust bearings. Instrumentation on the bearing pads measures the height of the oil film between the pad and the bearing runner. Techniques to analyze the film height record are developed and discussed. The analysis techniques present the unwieldy data in a compact form for assessment of bearing condition. The techniques are illustrated by analysis of a small sample of film height records from each of the three 64-m antennas. The results show the general condition of the bearings at DSS 43 and DSS 63 as good to excellent, and at DSS 14 as marginal.

I. Introduction

This report describes analyses of film height records of the DSN 64-m antenna azimuth thrust bearings. The 64-m antennas use an oil film hydrostatic azimuth bearing to support the 6.5-million-pound rotating weight of the alidade and tipping structure. The bearing configuration has a large oil pad at each apex of the triangular-shaped alidade base. Figure 1 shows the antenna and the azimuth bearing arrangement. The bearing pads have instrumentation to measure and record the height of the oil film between the pad and the supporting runner during operation of the antenna.

Film height records from each of the three antennas have been analyzed. Each record was analyzed in two ways: (1) calculation of the mean, standard deviation about the mean, and population histogram of the film height; (2) calculation and examination of the film height versus angular frequency spectrum. Also, averaged values of film height parameters versus azimuth angle have been developed for DSS 14. The analyses, which greatly compress the data on the records, imply the present and previous health of the bearings: good to excellent at DSS 43 (Australia) and DSS 63 (Spain); marginal at DSS 14 (Goldstone). The film height spectra clearly show the effects of the 11 regularly spaced bearing runner segments and provide quantitative assessment of the film height fluctuations resulting from the runner joints.

The analysis techniques used are simple in concept but laborious to apply with present equipment. The DSN plans to install equipment to produce digital film height records directly and to provide modest on-site data reduction capability. The equipment will be installed at DSS 14 in May 1983,

and at DSS 43 and DSS 63 in the fall of 1983. When that has been done, the techniques will be inexpensive and practical to apply and will be useful for support of operations and for maintenance planning of the bearings.

In the following, we present in order: the methods of taking and preparing the bearing film height data records, the techniques for data analysis, and the results of the analyses of the three antenna bearings. We close with suggestions on further development and application of the film height analysis techniques.

II. The Method of Taking and Preparing Film Height Records

A. Taking the Records

Each of the three bearing pads has precision linear transducers mounted to its periphery. The transducers are set to zero when the pad is sitting on the bearing runner. When the bearing is floated by oil supplied to recesses in the pads, the transducers measure the height, at their location, of the pad above the runner surface. The height is typically 0.005 to 0.010 in. for a properly operating bearing. The most heavily loaded pad, the rear pad (designated pad 3), has eight transducers as shown in Fig. 2; the other two pads have a transducer at each corner.

Film height records are customarily made from pad 3. Records are made several times each year. The transducers are connected to provide a voltage proportional to displacement; the voltage moves a stylus on a multichannel strip chart recorder. A normal recording has a calibrated trace from each of the eight transducers on pad 3. The azimuth position of the antenna is indicated by pips on the chart at 1° intervals as the antenna rotates, and the chart paper runs, at approximately constant speeds. A data run consists of a transducer zero calibration, a 360° rotation of the antenna, and a final transducer zero calibration. The run produces a chart 1/2 meter wide and 6 meters long. It is dated and annotated. The instrumentation is evidently accurate and stable within ≈0.0002 in.

B. Preparing the Data for Analysis

The traces were read and recorded by hand. Normally, five film height traces were read – from the transducers at the four corners and at the right center of the pad. They were read at 1° azimuth increments, to an accuracy of ≈±0.0002 in. of film height. The process takes about 20 hours per record; 11 records were read – four each from DSS 14 and DSS 43, and three from DSS 63.

III. The Methods of Data Reduction and Analysis

Correlation and spectrum analysis techniques have been used for solving problems and analyzing data in many fields of engineering. Here we calculate the autocorrelation function of the film height versus azimuth data and then obtain the power spectrum of the film height by Fourier transform. That is a standard procedure.

A conventional electrical waveform is voltage as a function of time and the related power spectrum is in units of watts per unit of frequency. In our case, the coordinates are film height as a function of azimuth rotation. Thus, the unit of spectrum power is film height squared, and the frequency is cycles of film height variation per azimuth revolution. The purpose in obtaining the spectrum of the film height is to provide for convenient examination of periodicities and other frequency characteristics of the data.

Histograms of film heights versus frequency of occurrence and the mean and the standard deviation about the mean of the film height are calculated. These, too, are standard calculations.

The power spectrum contains no phase information, so it cannot identify a characteristic associated with a particular azimuth angle. To do that, we return to the film height versus azimuth angle data and calculate azimuth-dependent film height parameters.

The details of the calculations and some sample results are presented in the following paragraphs.

A. Calculation of the Autocorrelation Function of Film Height

The autocorrelation function $C_{ff}(\gamma)$ is defined by:

$$C_{ff}(\gamma) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(\theta) f(\theta+\gamma)\, d\theta \qquad (1)$$

which is reformulated for our computational purpose as:

$$C_{ff}(N\Delta\theta) = \frac{1}{360} \sum_{\theta=1}^{360} f(\theta) f(\theta + N\Delta\theta) \qquad (2)$$

where $\Delta\theta = 1°$, $N = 1, 2, \ldots, 360$, and $f(\theta)$ is film height ($\times 10^{-3}$ in.).

The set of data f is reduced to zero mean before using Eq. (2) for computation.

1. **The computer program.** The program used to calculate the autocorrelation function of Eq. (2) was written in BASIC for the Hewlett Packard HP-85 computer used for the computations; it is given as a flow chart in Appendix A. The program reads the data file, then creates a new file which stores the computed autocorrelation data. Notice that it is necessary to compute only 180 points instead of 360 points for Eq. (2) (i.e., $C_{ff}(N\Delta\theta)$ is symmetric in N about every integral multiple of 180°); however, the subprogram is consistent with Eq. (2). That means the program will compute 360 points; the resulting autocorrelation function will be mirror imaged about θ = 180°

2. **Sample results.** A sample result of the calculation of the autocorrelation function by the BASIC program implementation of Eq. (2) is given in Fig. 3. It is for the right-center probe data of DSS 43 taken in October 1978. $C_{ff}(N\Delta\theta)$ is plotted in terms of $N\Delta\theta$. According to Eq. (2), $N\Delta\theta$ is in units of degrees and $C_{ff}(N\Delta\theta)$ is in units of square inches ($\times 10^{-6}$).

B. Calculation of the Power Spectrum of Film Height

Applying the Fourier transformation to the autocorrelation function will produce the power spectrum function. The purpose of doing this is to achieve a function in the frequency domain. Consider the finite Fourier series:

$$f(\theta) = a_0 + \sum_{i=1}^{M} a_i \cos \frac{2\eta i\theta}{T} + b_i \sin \frac{2\eta i\theta}{T} \qquad (3)$$

where

$$a_0 = \frac{1}{T}\int_0^T f(\theta)\, d\theta$$

$$\left.\begin{aligned} a_i &= \frac{2}{T}\int_0^T f(\theta) \cos \frac{2\eta i\theta}{T}\, d\theta \\ b_i &= \frac{2}{T}\int_0^T f(\theta) \sin \frac{2\eta i\theta}{T}\, d\theta \end{aligned}\right\} \quad i = 1, \ldots, M.$$

In our case of application, $C_{ff}(N\Delta\theta)$ is $f(\theta)$ in the above a_i and b_i. Furthermore, T = 360° and M = 180°, and one can write:

$$a(i) = \frac{1}{180} \sum_{i=0}^{180} \sum_{j=1}^{360} C_{ff}(j) \cos(i*j) \qquad (4)$$

$$b(i) = \frac{1}{180} \sum_{i=0}^{180} \sum_{j=1}^{360} C_{ff}(j) \sin(i*j) \qquad (5)$$

Finally, the film height power spectrum function, the Fourier transform of the autocorrelation function, is given by:

$$C(i) = [a(i)^2 + b(i)^2]^{1/2} \qquad (6)$$

1. **The computer program.** Equations (4), (5), and (6) are implemented in the program flow charted in Appendix B. The main program will input the autocorrelation data file and create the spectrum function file which is used to store the computed data.

2. **Sample results.** A sample result of the calculation of the film height power spectrum is shown in Fig. 4. It is for the right-center probe data of DSS 43 taken in October 1978. It is the power spectrum derived from the autocorrelation function of Fig. 3. In Fig. 4, the horizontal axis is the frequency of film-height variation for a revolution around the runner. The vertical axis is film height squared ($\times 10^{-6}$ in.^{2}) per cycle per azimuthal revolution (CPR).

The integral of the Fig. 4 power spectrum over a full azimuth rotation (360°) is 1.92 $\times$ 10^{-6} in.^{2}, which is identical to the value of the autocorrelation function at N = 0 in Fig. 3. This is a formal identity displaying selfconsistency of the calculations of the two computer programs.

An eleven CPR fundamental with harmonics out to the 6th are prominent features of the spectrum in Fig. 4. There are eleven segments in the azimuth bearing runner.

C. Calculation of the Histogram, Mean, and Standard Deviation of the Film Height

The purpose of generating a histogram is to show the distribution of the film-height data. A broad distribution of the film-height data means a large variation in film height. The mean, $\bar{h}$, and standard deviation about the mean, σ, of the data are important parameters of the distribution of film heights.

1. **The computer program.** The computer program used to generate histograms is in the Hewlett Packard HP-85 Standard Pac software. The program is written in BASIC and consists of 315 lines. It will input a specified data file from the tape and automatically sort the values into intervals or bins of equal width. Then a histogram and cell statistics can be generated and printed out at the user's request.

Two changes have been made in the program to fit our application. The size of the declared array has been changed

to the size of the data set that is used to produce the histogram. Also, a few lines of codes have been added to the cell statistics subroutine to print out the mean and the standard deviation of the sample data.

For further information and a listing of the program, refer to the documentation of Histogram Generator in the Hewlett Packard HP-85 Standard Pac.

2. **Sample results.** Figure 5 is a sample histogram. It is for the right-center probe of DSS 43 taken in October 1978. The horizontal axis is film height in inches $\times 10^{-3}$. The left side of the vertical axis gives the numbers of observations, and the right side of the vertical axis gives the percent of relative frequency.

The printed output on Fig. 5 gives the mean $\bar{h}$ and the standard deviation σ as 10.70 and 1.38 $\times$ 10^{-3} in., respectively. The histogram of Fig. 5 resembles a Gaussian distribution. Let us examine that. The values $\bar{h} \pm \sigma$, 2σ, and 3σ are:

	σ	2σ	3σ
$\bar{h}-$	9.31	7.93	6.55
$\bar{h}+$	12.08	13.47	14.85

Comparing the above with the cell statistics, the approximate number of film height samples (i) outside of the 1σ, 2σ, and 3σ about $\bar{h}$ is 118, 21, and 3, respectively. Thus, there is approximately 32.8 percent, 5.8 percent, and 0.8 percent of the samples outside of $\bar{h} \pm \sigma$, $\pm 2\sigma$, and $\pm 3\sigma$, respectively. For a true Gaussian distribution, the corresponding percentages are 31.7 percent, 4.6 percent, and 0.3 percent. That shows the distribution is, indeed, nearly Gaussian in character.

Finally, the value of the variance, σ^2, is 1.92 $\times$ 10^{-6} in.^{2}, which is identical to the value of the power spectrum integration and the value of the autocorrelation function at $N = 0$, as it must be. This again shows correctness of the procedure including the computer programs used.

D. Calculation and Plots of Compressed Film Height Parameters Versus Azimuth Angle

In addition to the spectral and statistical analyses, we need techniques for displaying and examining film height parameters related to a specific bearing azimuth angle. The display of the data must be compact, i.e., reduced from the 1/2 X 6 meter original record to a format perhaps the size of this piece of paper. To do that, we proceed as follows.

Consider the equation:

$$\bar{h}(\theta, W) = \frac{1}{W+1} \sum_{\theta_W = \theta - \frac{W}{2}}^{\theta_W = \theta + \frac{W}{2}} f(\theta_W) \qquad (7)$$

where θ_W is an integer; $\bar{h}(\theta, W)$ is the mean film height at azimuth θ, smoothed over the window W centered on θ; $f(\theta_W)$ is the film height data at azimuth θ_W. The smoothed film height $\bar{h}(\theta, W)$ from Eq. (7) can be displayed in a compact graphical form.

In the same way, other smoothed functions of the film height versus azimuth can be computed. e.g., $\sigma(\theta, W)$ and $\bar{h}/\sigma\ (\theta, W)$. These techniques are used to develop plots in Section IV.C.1 of this article.

IV. Results of Reduction and Analysis of Film Height Records From the 64-Meter Antennas

A sampling of film height records from the three antennas was reduced and analyzed using the techniques described in the preceding sections. The records are identified in Table 1.

Spectrum plots are shown of the film height power versus azimuthal angular frequency for the *right center*[1] probe data of each record. Film height statistics from the *four corner* probes of each record are given. Also, locally averaged azimuth-dependent film height parameters are calculated from two of the DSS 14 records.

The existing analytical design model of the bearing does not predict significant detailed behavior shown by analyses of the experimental film height data. Therefore, we give physical interpretations of some spectral and statistical results to help in future model development.

We first discuss the *set* of spectral and film height statistics data. Then, because each antenna bearing has a different film height signature, the interpretations of their data are discussed individually.

A. The Set of Film Height Power Spectra

The film height power spectra of the right center probe of pad 3 for the several records are displayed in Figs. C-1 through C-11 of Appendix C.

[1]Spectra of four corner probe data were also computed and analyzed; for brevity, that work is omitted from this article.

1. **Evidence of runner joint effects.** In all of the spectra, but to varying degree, there is evident an 11-CPR component with its higher harmonics (i.e., 22, 33, etc., CPR). The 11 joints, the only regular inhomogeneity in the runner, produce the fluctuation power (variance, σ^2, in film height) at the 11-CPR and harmonically related frequencies.

2. **Significance of runner joint effects.** We can illustrate the relative significance of the bearing segment joints in the total variance of the bearing film height. Figures 6(a), (b), and (c) show plots of the total film height variance, which is the integral of the power spectra over the full range of azimuthal frequencies, excepting zero. Those figures also show the variance due to the power contribution at 11, 22, 33, etc., CPR. The relative contribution of the bearing joints to the total film height variance is quite different for each antenna; the differences are qualitatively evident from the spectra of Appendix C.

Further discussion of the spectrum plots of the individual antennas will be given in Section IV.C.

B. The Set of Film Height Statistics

Histograms, and values of the mean, and the standard deviation about the mean of the film height data were computed for the set of records given in Table 1. (The DSS 43 record of March 1982 was not included because, at that particular time, one of the four corner probes was defunct.) The results were in the format shown in Fig. 5.

1. **Comparison of statistics with Gaussian model.** In Section III.C.2 on Analysis Methods, we observed that the distribution of film height data is Gaussian-like. Table 2 addresses that proposition for the complete data set of 14,400 points; it is derived from the calculation of mean heights and standard deviations as shown in Fig. 5.

Referring to Table 2, the percentage of samples within the mean ±1, 2, and 3 standard deviations is generally close to that of a Gaussian distribution. The table also shows the *actual* percentage of the 1440 film-height samples in each record which are *below* the value $\bar{h} - 3\sigma$. Note that, with the progress of time, the number of very low film heights is decreased, and the "wings" of the distribution are skewed – presumedly a result of maintenance shimming of the bearing to correct specific low film height areas.

2. **Condition of a bearing inferred from statistics.** The purpose in displaying the Gaussian-like distribution of the film height data is to develop a simple technique to show the general condition of the bearing from a film height record. That technique is displayed in Figs. 7(a), (b), and (c), which are plots of $\bar{h}$, and $\bar{h} \pm 3\sigma$ for the three bearings. Because of the approximate Gaussian distribution, there should be very few (~2) points with film height lower than $\bar{h} - 3\sigma$. Also shown on the figures are the *actual* lowest film heights taken from the records and the design, minimum normal operating, and minimum safe film heights in the bearing.

It is evident that analysis of a more complete set of historical records, e.g., annually, would provide a basis for a useful extrapolation into the future of the general condition of a bearing.

C. Discussion of Individual Antennas

As noted, the bearings of the three antennas have different individual film height characteristics. The interpretations of their film height spectra and statistics are discussed individually in the following paragraphs.

1. DSS 14 characteristics

a. Runner joint (11 CPR) components in spectra. The first two spectra (August 1971, July 1981; Figs. C-1 and C-2) clearly show power at the 11 CPR frequency of the bearing segments. In the two later spectra, Figs. C-3 and C-4, the 11 CPR components are not as clear although there is considerable power in the general regions around 11 CPR and related harmonics. That behavior is consistent with the variance plot of Fig. 6(a), which shows that the relative contribution of the 11 CPR and exact harmonic components in the two latest records (December 1981 and February 1982) is down by a factor of about 2 from the previous record (July 1981). It may be a coincidence, but several gaps of ~0.010 to 0.040 in. at the bearing joints were closed during bearing maintenance in November 1981.

b. Broadband components in spectra. There is broadband noise power out to 60-70 CPR in the film height spectra of all antennas. It is especially prominent in the DSS 14 spectra. We consider a model for the generator of that noise. The bearing pad is very stiff (20 × 40 × 60 in. solid steel); we would expect that faulty support of the bearing runner would produce film height variations that are correlated over an angular extent related to the pad dimension along the runner. A possible mechanism is depicted in Fig. 8.

The actual film height records show many randomly spaced fluctuations which resemble the little cosine wavelet of the sketch, both in shape and angular dimension.

The power spectrum shape of an idealized cosine wavelet one pad-length long is as shown in Curve A of Fig. 9. It is superimposed on the February 1982, DSS 14 spectrum of Fig. C-4. The shape of the wavelet spectrum resembles the envelope of the film height spectrum above about 20 CPR.

For comparison, the spectral shapes are shown for film height wavelets which are one-half and twice the length of the pad (Curves B and C, respectively).

c. Low frequency components in spectra. Another distinguishing characteristic of the DSS 14 spectra is the large peaks in the 1-3 CPR range. One might suspect that the spectral analysis process is producing artifacts for frequencies so close to zero. To examine that, we return to the original film height records. Figure 10 shows the actual right center probe film height versus trough azimuth for the August 1971 DSS 14 record. In that figure, the one sample per degree data have been smoothed over a 20° wide azimuthal window. Otherwise, the data are as taken from the original record. The record has a readily discernible 1 CPR variation, qualitatively consistent with the spectrum plot. To further illustrate that consistency, we proceed as follows. The power in the prominent 1 CPR component of the August 1971 record (Fig. C-1) is $\approx 0.26 \times 10^{-6}$ in.^{2}. The amplitude of a 1-CPR sine wave of that power is:

$$A = \sqrt{2P} = \sqrt{2 \times 0.26 \times 10^{-6} \text{ in.}^2} = 0.72 \times 10^{-3} \text{ in.}$$

An approximate fit of a 1-CPR sine wave of that amplitude added to the average film height $\bar{h}$ is depicted in Fig. 10. It is qualitatively consistent with the original smoothed data. The possible cause of the very slow variations in film height with azimuth is not evident; the variations do, however, exist.

d. Search for the evidence of grout keyways in spectra. When this work was started, a principal purpose was to identify the effects of suspected faulty grout keyways under the DSS 14 bearing runner. There are 40 regularly spaced keyways under the bearing runner. They are formed by rectangular cavities in the top of the concrete pedestal. When the grout was originally placed between the runner and the pedestal, it was supposed to fill the cavities. It is known from coring samples that the filling was not uniformly sound. The purpose of the keyways is to lock the grout securely to the pedestal. We anticipated identifying the 40-CPR signatures in the film height spectra.

We did not find evidence of the keyways in any of the right center probe or four corner probe spectra. To further investigate the keyway matter, pad 3 was fitted with additional probes at the middle of its long dimension, see Fig. 11. It was expected that the probes at the centers of the long dimension of the pad would be most effective in detecting poor support of the runner above the 40-grout keyways. The February 1982 record has data from the new CF and CR probes. The spectra from that data *do not* show significant 40-CPR power (the spectral plots are not included here). This result, and the fact that the film height spectra of the four corner and right center probes also showed no 40-CPR power, leads us to the unexpected conclusion that the grout keyways are not significant systematic contributors to degraded performance of the DSS 14 bearing.

e. Investigating bearing support using new pad center probes. If the support under the bearing pad is compliant relative to the pad, the bearing runner will be most depressed in the middle of its long dimension. We should measure that effect by differencing the film height at a center probe and the average of the film heights at the two ends of the same face of the pad. This is, referring to Fig. 11:

$$\left(h_{CF} - \frac{h_{LF} + h_{RF}}{2}\right); \qquad \left(h_{CR} - \frac{h_{LR} + h_{RR}}{2}\right)$$

Plots of those differences versus trough azimuth, smoothed over ±10° in azimuth, are shown in Figs. 12 and 13. From those figures one is led to suspect that the support is better under the inner (rear) area of the bearing runner than it is at the outer (front) area, and that the region around 170° of trough azimuth is especially well supported. Only a single record with the new CF and CR probes has been analyzed; the results need further validation to ensure that they are authentic.

f. Right center versus corner probe film heights, DSS 14 compared with overseas antennas. The DSS 14 bearing has another special characteristic. Table 3 provides a comparison of the mean film height at the right center probe with the average of the four corner probes for the records of all antennas.

The DSS 43 and 63 bearings have equal or slightly greater mean film heights at the center probe compared to that at the four corner probes. All but the earliest record from DSS 14 show lower film heights from the right center probe. This may be a result of the extensive use of radially contoured shims under the bearing runner at DSS 14.

g. An examination of locally averaged film height and standard deviation. A high average film height $\bar{h}$ with a low standard deviation σ are the desired characteristics of a bearing. In Fig. 14, those parameters from the February 1982 DSS 14 record are plotted versus trough azimuth. They have been smoothed over ±10° in azimuth. Also, the quotient, $\bar{h}/\sigma\ (\theta)$, is plotted – the higher the $\bar{h}/\sigma\ (\theta)$ quotient, the better the bearing in the ±10° region around the associated azimuth value. The regions of trough azimuth over which a portion of the pad is over a runner joint are indicated in Fig. 14.

The presentation of Fig. 14 displays significant characteristics of the DSS 14 bearing. The average smoothed film

height, $h(\theta)$, is decent ($9.2 \pm 0.7 \times 10^{-3}$ in.). There are some rather large areas between the joints where the $\sigma(\theta)$ is good ($\sigma < 1.7 \times 10^{-3}$ in.). The poor areas are clearly shown by the $\bar{h}/\sigma$ (θ) plot; they are systematically associated with the runner joints. Finally, all of the plots are distinctly mirror-imaged about a trough azimuth of 162°. The mirror imaging is especially evident in the $\bar{h}/\sigma$ (θ) plot; the cause is at present a mystery.

2. DSS 43 characteristics

a. Runner joint (11-CPR) components in spectra. The dominant features of the DSS 43 spectrum plots (Figs. C-5, 6, 7, and 8) are the 11-CPR component and its harmonics. As presented in Fig. 6(b), a large fraction of the total film height fluctuation power is contained in those components.

Plots of the power in the 11-CPR components taken from the earliest (September 1976) and the latest (March 1982) record spectra are depicted in Fig. 15. The characters of the two other record spectra (October 1978 and June 1980) generally lie between the two shown. The basic shape of the spectrum has not changed substantially in the 6 years between the records, but the power in each component has increased considerably; e.g., the 11-CPR fundamental ($n = 1$ in Fig. 15) has almost doubled in power. Evidently some changes have occurred over the years in support under the runner joints or the continuity of the joints, or both.

b. Model of joint-dominated film height fluctuations, DSS 43. By studying the original film height records and doing some trial-and-error spectrum shape fitting, we synthesized the very simplified film height wave shape shown in Fig. 15. The value of $\theta_1 = 3.74°$ in the model is the angular half width of the bearing pad. The 32.7° long pattern repeats 11 times in one azimuth revolution, corresponding to the number of runner joints. The ratio $a_2/a_1 = 0.3$ is representative of the average behavior of the experimental film height data in the vicinity of the joints. The amplitude $a_1 = 2.1 \times 10^{-3}$ of the model waveform was used to put its spectrum between the two experimental spectra. To closely match the September 1976 record, select $a = 1.7 \times 10^{-3}$ in.; to match the March 1982 record select $a_1 = 2.4 \times 10^{-3}$ in. The actual experimental film height data do not have the sharp corners of the model wave form but rather the smooth contours of the cosine wavelet. However, that difference only modestly affects the higher order component powers, not the dominant lower order ($n = 1, 2$) components; the simple model was used for ease of calculation.

The model displays a persistent lower film height in the vicinity of the runner joints; that is what produces the large 11-CPR fundamental frequency component in the spectra. The spectra show that the magnitude of this effect has increased with time. The physical cause, to our knowledge, is not now known. Although the bearing appears to be in good condition, this characteristic should be understood in the context of its long-term maintenance. Abrupt steps at the runner joints are not much in evidence at DSS 43, either in the raw data or in the derived spectra. Therefore, they were not included in the model. One should remember that the fluctuation model waveform (cf Fig. 15) "sits" upon a healthy mean film height of 9 to 10 thousandths of an inch.

c. Broadband components in spectra. In Fig. C-8, the DSS 43 broadband noise can be compared with the spectral envelope of 1- and 2-pad-length-long cosine wavelets. The shape of the bearing noise spectrum appears to lie between those of the two models.

3. DSS 63 characteristics

a. Runner joint (11-CPR) components in spectra. The DSS 63 spectrum plots (Figs. 3-9, -10, and -11) show power at the 11-CPR fundamental and its harmonics. The power is distributed to the higher harmonics more so than in the DSS 43 spectra; it is visible out to the 8th harmonic (88 CPR).

Plots of the power in the 11-CPR components from the spectra of the earliest (July 1973) and the latest (March 1982) records are depicted in Fig. 16. The character of the other record spectrum (June 1978) generally lies between the two shown. The shape of the spectrum has changed noticeably, and its magnitude, while still relatively small, has increased in the 8-1/2 years between the records. By examination of film height spectra at one- to two-year intervals in the future, any significant progress of deterioration of the joint support or continuity can be identified.

b. Model of joint-dominated film height fluctuations, DSS 63. Again by studying the original film height records and doing trial-and-error spectrum shape fitting, we prepared a simple model for the periodic film height waveform (cf Fig. 16). The parameters of the model were selected so that its spectrum envelope lies within that of the two data records shown. It is very similar to that of the June 1978 record, not shown. Increasing τ (cf Fig. 16) brings the model toward the March 1982 record spectrum; decreasing τ brings it toward the July 1973 record spectrum. This relates to a slow degradation of pad support in the joint regions versus the support at the regions midway between the joints.

The DSS 63 model has a relatively modest amplitude. The humped valley around the joints is broad with an average height not far below that of the plateau between the joints. The narrow downward triangular spikes at the joints mimic abrupt 0.001- to 0.002-in. steps displayed at most of the

joints in the DSS 63 film height records. These rapid variations produce the higher harmonics ($n = 6, 7, 8$) in the DSS 63 spectra.

c. Narrowband components (besides 11 CPR) in spectra. The DSS 63 spectra (Figs. C-9, 10, and 11) consistently show power at 9 CPR, 38 CPR, and 90 CPR. Added together, those components contain 6 percent, 17 percent, and 15 percent of the total fluctuation power, σ^2, in the July 1973, June 1978, and March 1982 spectra, respectively.

The 9-CPR component in the June 1978 record is the largest of the set at 0.14×10^{-6} in.^{2}. That corresponds to a 9-CPR sine wave of peak-to-peak amplitude $2a = 2\sqrt{2P} = 2\sqrt{2 \times 0.14 \times 10^{-6}\ \text{in.}^2} \cong 1.1 \times 10^{-3}$ in., which is neither negligible nor really significant. We have no postulation on the source of the 9, 38, or 90-CPR spectral components.

d. Broadband components in spectra. In Fig. C-11, the DSS 63 broadband noise and the 1-pad-length cosine wavelet spectral shapes are compared. There is qualitative agreement, indicating consistency with a model of the type depicted in Fig. 8.

V. Discussion

Our conclusions from analyses of the several bearing records is that the spectral and statistical analyses techniques can provide reliable information on the previous and present conditions of a bearing. Because the results appear consistent, not erratic, over long time spans, useful extrapolation to future bearing condition should be available using the methods. The spectral analysis and azimuth-related film height parameter analysis also provide systematic pointers to guide investigation of specific existing or incipient faults in the bearings.

The methods get the voluminous data from the film height records into a form which can be digested and preserved and can be usable for comparisons over long time spans. Fundamentally, the results are easy to obtain and, with familiarity, easy to interpret.

Analysis of additional historical records, and those currently being obtained, requires the laborious reading of the analog strip chart traces. However, that is just painful, not really very expensive or time consuming. Within a year, when the current plans are carried through, the analog recording equipment will be replaced with digital equipment. The digital records then obtained will be in a form to directly input the analysis programs we have used.

A. Regarding the Statistical Analyses

The mean film height $\bar{h}$ and standard deviation about $\bar{h}$, σ, of the four corner probes appear to be meaningful measures of the general condition of a bearing. If a reasonable number of the historical records (e.g., two or three per year to reduce environmental effects) were analyzed, we expect that a useful extrapolation of future bearing maintenance needs would result. The calculations and interpretation of the results are especially straightforward.

B. Regarding the Spectral Analyses

The spectra derived faithfully display the systematic nature of the film height fluctuations and in a way that cannot be gleaned from studying the noncompressed raw data. The spectral analysis process can be useful in developing improved analytical models of the actual bearing performance (models which produce the experimentally observed behavior), and in assessing the results of maintenance work on the bearings.

A particular extension of the spectral analyses, which we expect would contribute to understanding the bearing operation, is the calculation of the *cross*-correlation and *cross*-spectra of the film heights at the two ends of the pad. In the model we discussed in the article to produce the cosine wavelet fluctuation from the right end probe, there should be produced simultaneously in azimuth a companion wavelet from the left end probe. The cross-correlation/cross-spectrum will quantitatively display that characteristic and thus support (or invalidate, or revise) the model. The raw data suggest the wavelets from the two pad ends do indeed occur at times simultaneously, as predicted by the model. This work is being planned.

C. Regarding the Calculation of Compressed Film Height Parameters Versus Azimuth Angle

With the data prepared for the statistical and spectral analyses, it can be easily manipulated to display in compressed form various film height parameters as a function of trough azimuth. We illustrated that, for example, with plots of locally averaged $\bar{h}(\theta)$, $\sigma(\theta)$, and the quotient $\bar{h}/\sigma(\theta)$. In that form, the film height data can be readily compared with other measured or calculated azimuth-related parameters of the bearing installation, e.g., with pedestal concrete modulus, to identify correlations.

In particular, we suggest that additional records of the new center front and rear probes at DSS 14 be analyzed. That will improve understanding of the bearing runner support as a function of azimuth and also across the width of the bearing (radially).

Reference

1. *The NASA/JPL 64-Meter-Diameter Antenna at Goldstone, California: Project Report*, Technical Memorandum 33-671, Jet Propulsion Laboratory, Pasadena, Calif., July 15, 1974.

Acknowledgments

I. D. Wells and W. H. Kuehn provided the records we analyzed. F. D. McLaughlin provided the antenna and bearing sketches. Discussions with M. S. Katow, H. D. McGuinness, F. D. McLaughlin, H. P. Phillips, A. A. Riewe, and I. D. Wells have been helpful in understanding engineering features of the bearing and its instrumentation.

Table 1. Film height records analyzed

(Each record is from the antenna rear pad, no. 3)

Antenna	Date recorded	Comment
DSS 14	08/27/71	Earliest record available
DSS 14	07/05/81	Prior to bearing fault of 9/81
DSS 14	12/02/81	After bearing fault rework
DSS 14	02/04/82	Includes probes at front and rear pad centers
DSS 43	09/21/76	
DSS 43	05/10/78	
DSS 43	06/11/80	
DSS 43	03/31/82	Only 3 of 4 corner probes available
DSS 63	07/25/73	
DSS 63	06/14/78	CW Az rotation (all above use CCW)
DSS 63	03/06/82	CW Az rotation

Table 2. Distribution of film heights at four corner probes, comparison with Gaussian distribution

Data	Percentage of samples within $\bar{h} \pm \sigma$	$\bar{h} \pm 2\sigma$	$\bar{h} \pm 3\sigma$	Percentage of samples Below $\bar{h} - 3\sigma$
Gaussian model	68.3	95.4	99.7	0.13
DSS 14 records:				
August 1971	68.2	94.9	99.3	0.56
July 1981	64.9	94.0	99.5	0.06
December 1981	64.3	96.2	99.7	0.08
February 1982	65.7	96.1	99.6	0.00
DSS 43 records:				
September 1976	64.2	93.5	99.7	0.25
October 1978	64.1	95.3	99.9	0.00
June 1980	64.8	96.5	99.9	0.00
DSS 63 records:				
July 1973	67.8	96.5	99.6	0.34
June 1978	65.9	94.3	99.4	0.21
March 1982	66.9	95.7	99.6	0.05
Average, all records (14,400 samples)	65.7	95.3	99.6	0.15

Table 3. Comparison of mean film heights of right center and four corner probes ($\bar{h}$ in in. $\times 10^{-3}$)

Record	$\bar{h}$ Right Center	$\bar{h}$ Four Corners
DSS 14, August 1971	10.4	8.2
DSS 14, July 1981	6.4	7.3
DSS 14, December 1981	7.5	9.4
DSS 14, February 1982	7.6	9.2
DSS 43, September 1976	10.7	10.5
DSS 43, October 1978	10.7	10.3
DSS 43, June 1980	11.8	11.3
DSS 63, July 1973	11.2	10.5
DSS 63, June 1978	11.1	10.9
DSS 63, March 1982	11.2	11.2

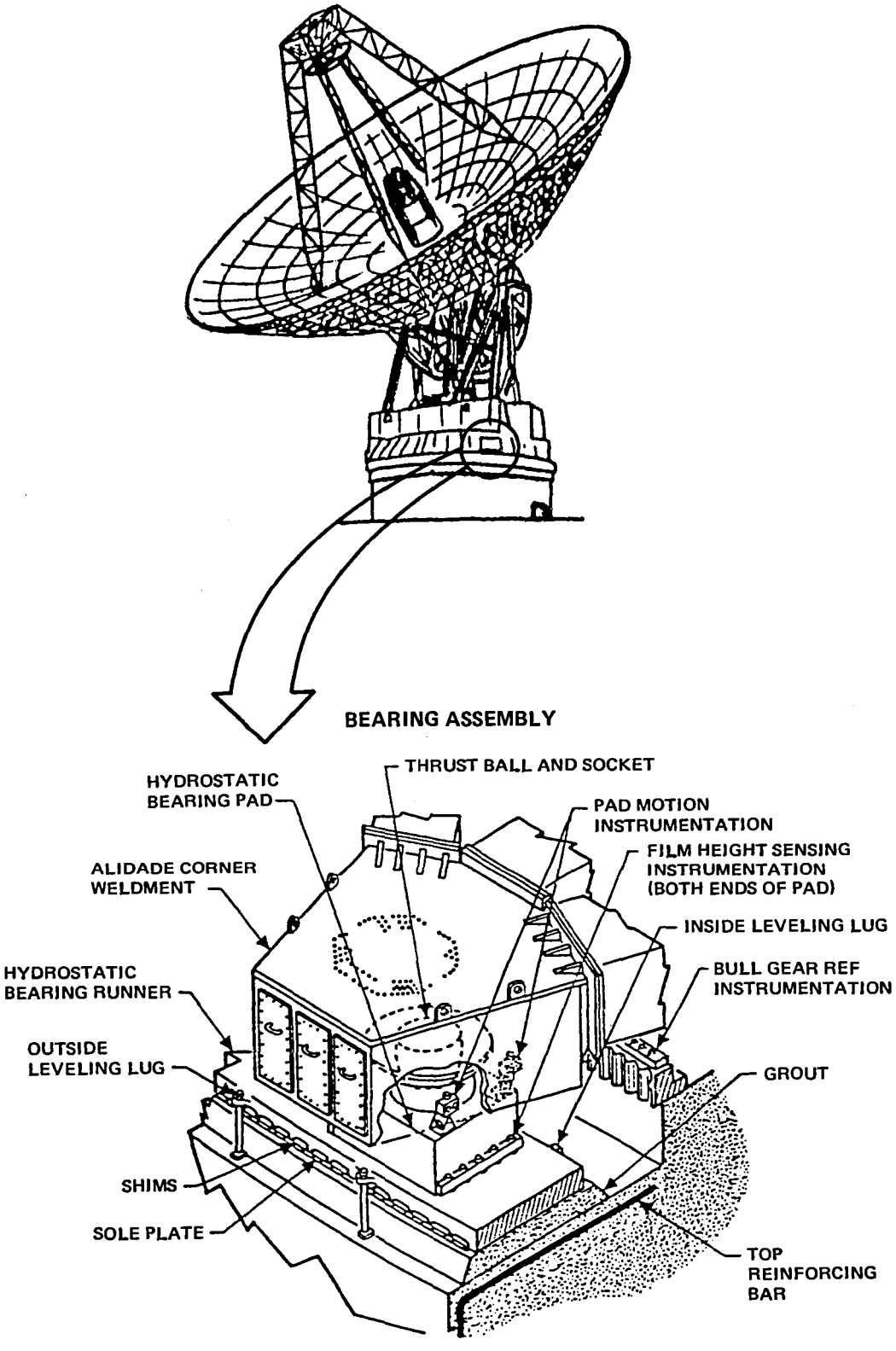


Fig. 1. 64-meter antenna showing azimuth thrust bearing arrangement

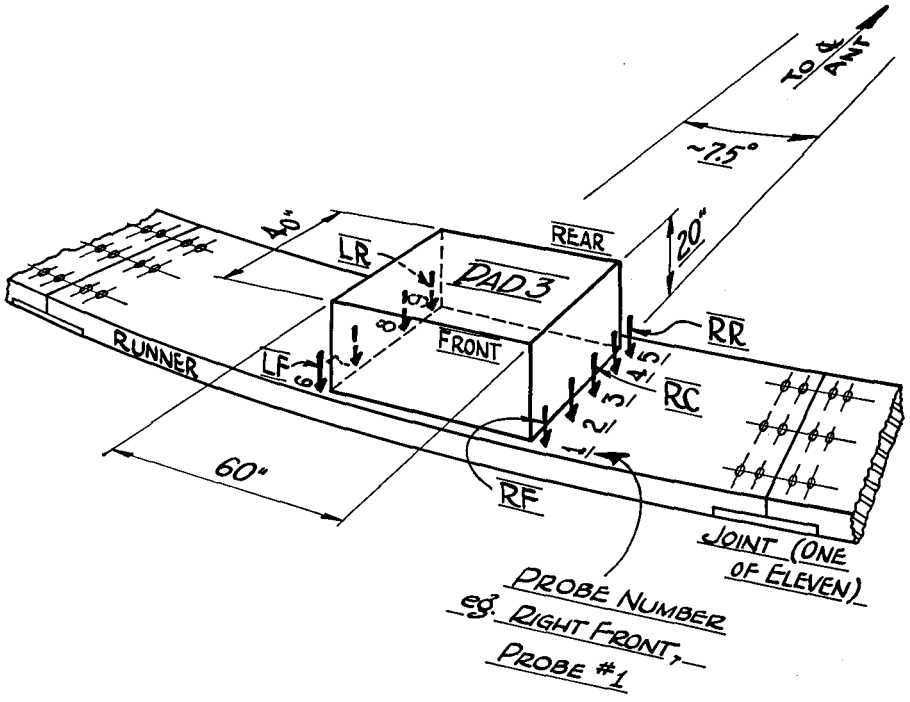


Fig. 2. Azimuth bearing pad with film height transducers

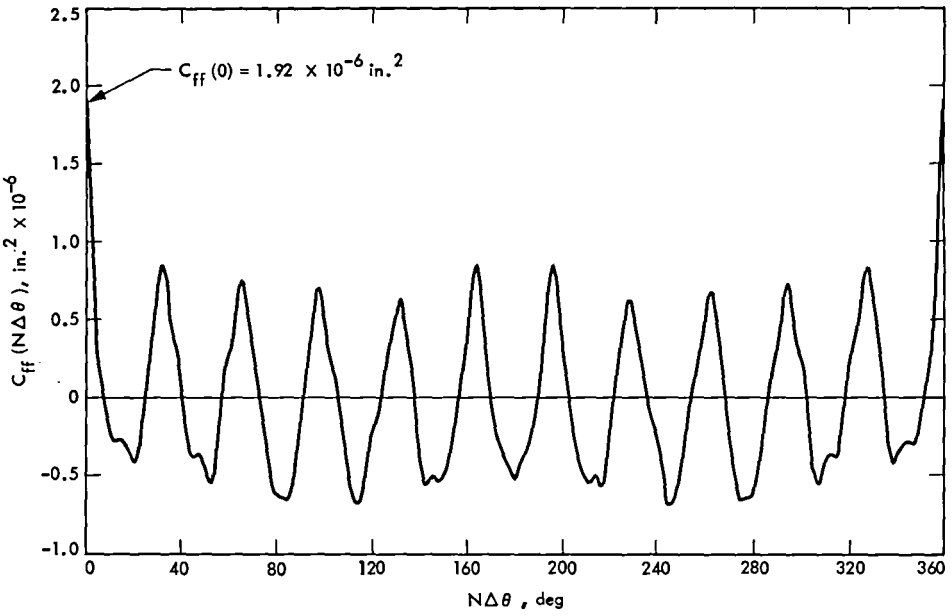


Fig. 3. Autocorrelation of film height; right center probe. DSS 43, October 1978

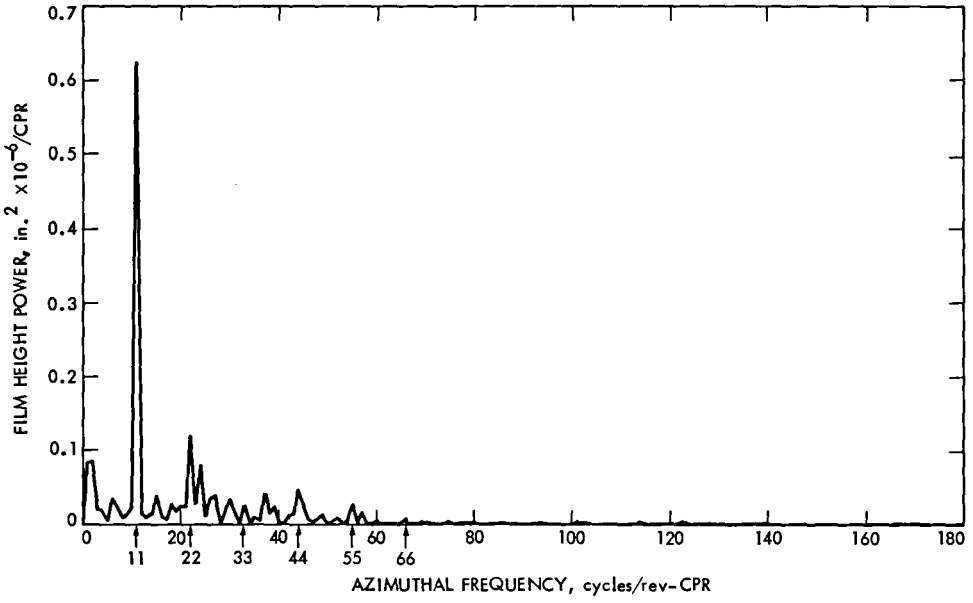
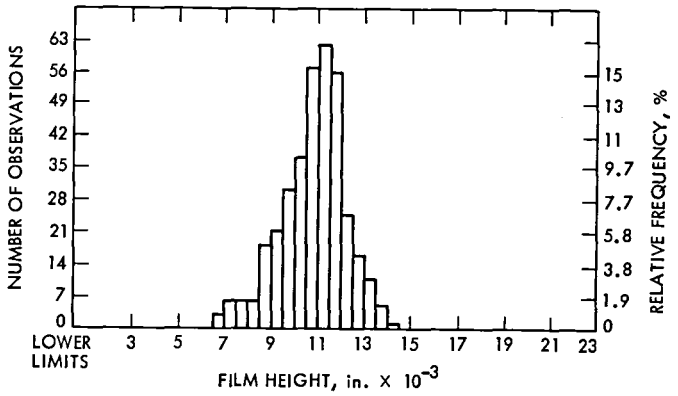


Fig. 4. Power spectrum of film height; right center probe. DSS 43, October 1978



CELL WIDTH = 0.5 X 10^{-3} in.

CELL STATISTICS

CELL #	LOWER LIMIT	NUMBER OF OBSERVATIONS	RELATIVE FREQUENCY
8	6.50	3	0.83
9	7.00	6	1.67
10	7.50	6	1.67
11	8.00	6	1.67
12	8.50	18	5.00
13	9.00	21	5.83
14	9.50	30	8.33
15	10.00	37	10.28
16	10.50	57	15.83
17	11.00	62	17.22
18	11.50	56	15.56
19	12.00	25	6.94
20	12.50	16	4.44
21	13.00	11	3.06
22	13.50	5	1.39
23	14.00	1	0.28

THE MEAN OF THE SAMPLE, $\bar{h}$, IS:
10.7 X 10^{-3} in.

THE STANDARD DEVIATION ABOUT THE MEAN OF THE SAMPLE, σ, IS:
1.38 X 10^{-3} in.

Fig. 5. Histogram and cell statistics of right center probe film height. DSS 43, October 1978

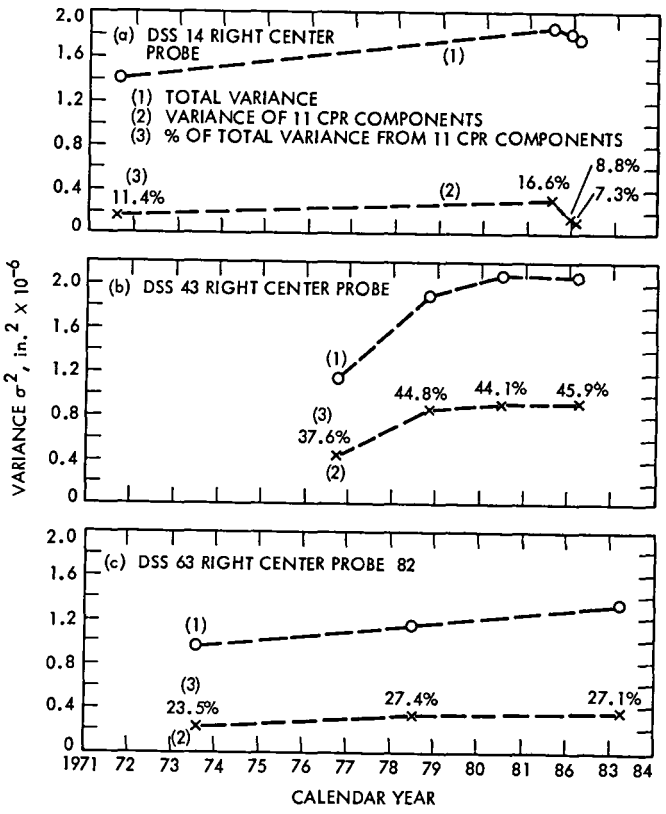


Fig. 6. Total variance and variance from 11-CPR components of film height

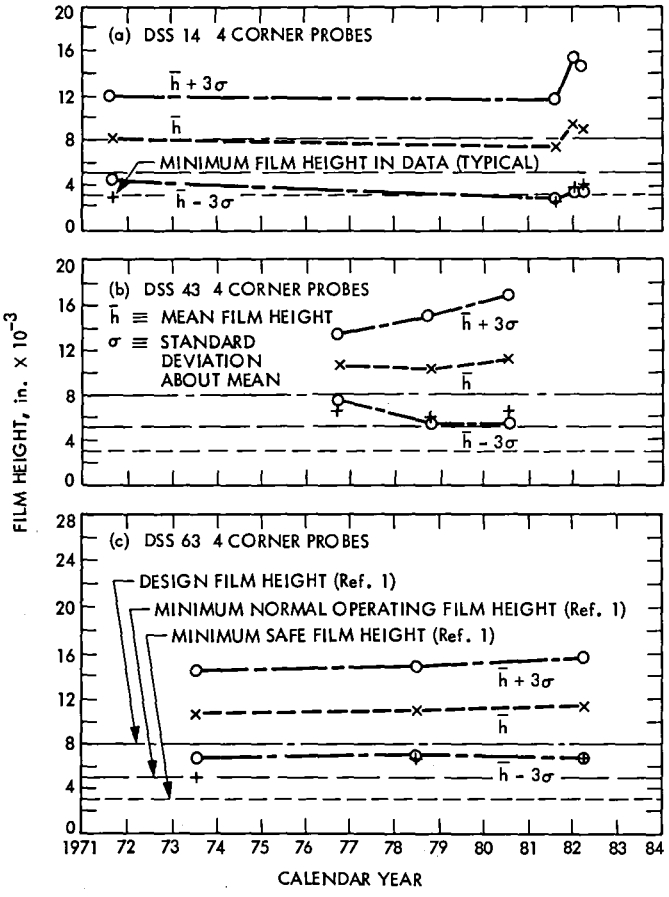
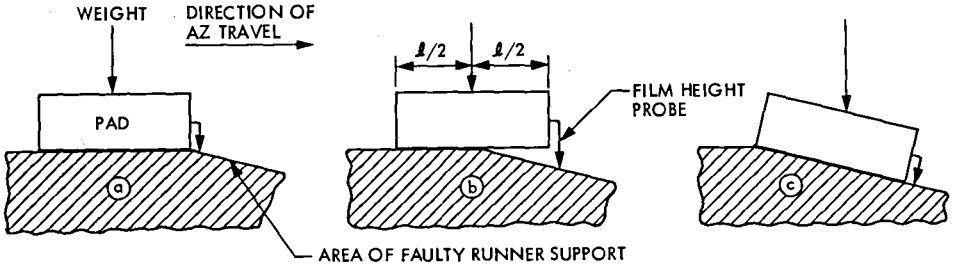


Fig. 7. Mean and mean ± three standard deviations of film height



FRONT OF PAD MOVES OVER START OF SLOPE IN ⓐ AND ITS FILM HEIGHT INCREASES, ⓑ. AFTER LOAD CENTER MOVES OVER SLOPE EDGE. PAD SETTLES DOWN, ⓒ. THE EXPECTED FILM HEIGHT VARIATION AT THE LEADING EDGE PROBE IS DEPICTED IN ⓓ BELOW.

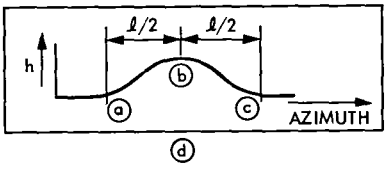


Fig. 8. Model of film height variation related to pad azimuth dimension

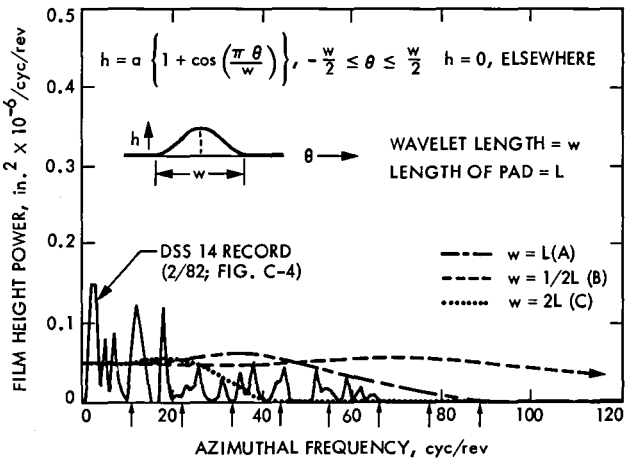


Fig. 9. Power spectra of model wavelets (with February 1982 DSS 14 spectrum)

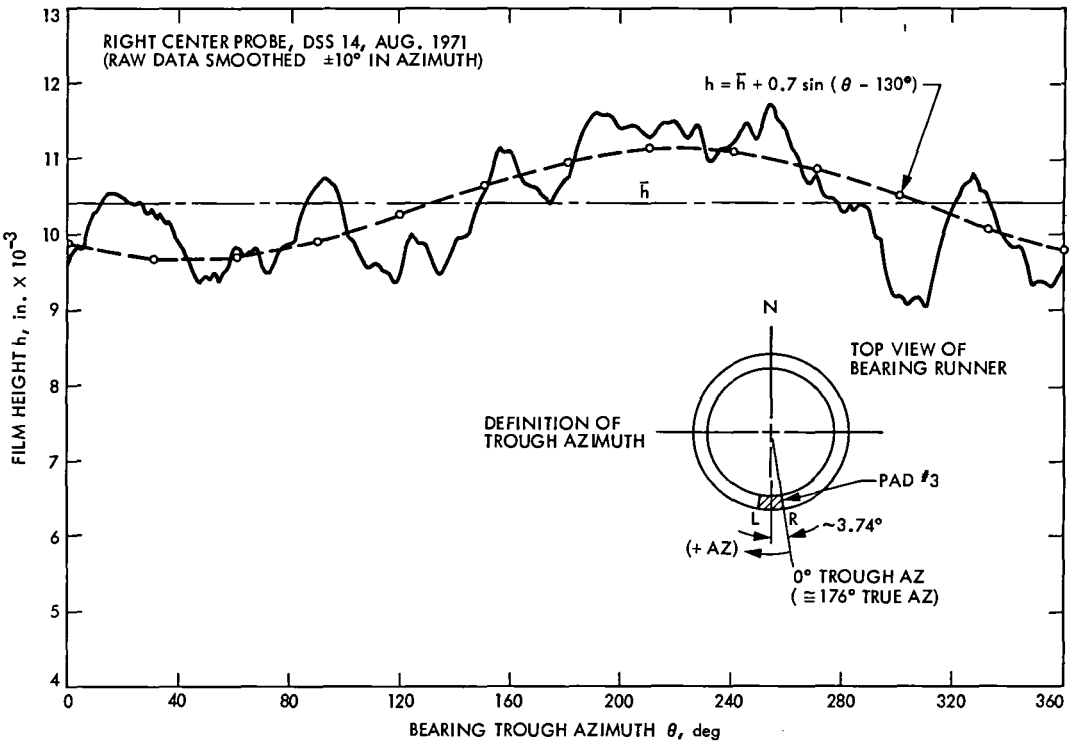


Fig. 10. Smoothed film height; approximate 2-CPR component

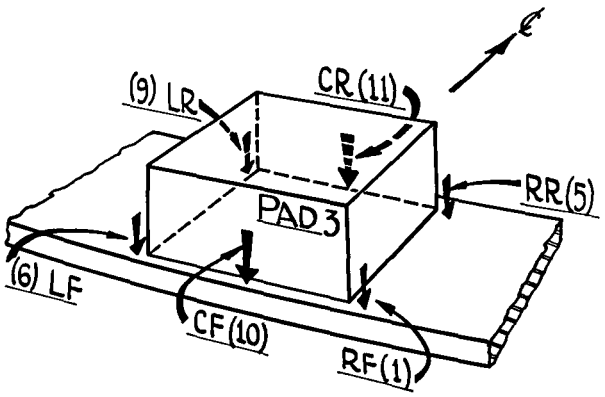


Fig. 11. Location of new center front (CF) and center rear (CR) film height probes on Pad 3, DSS 14

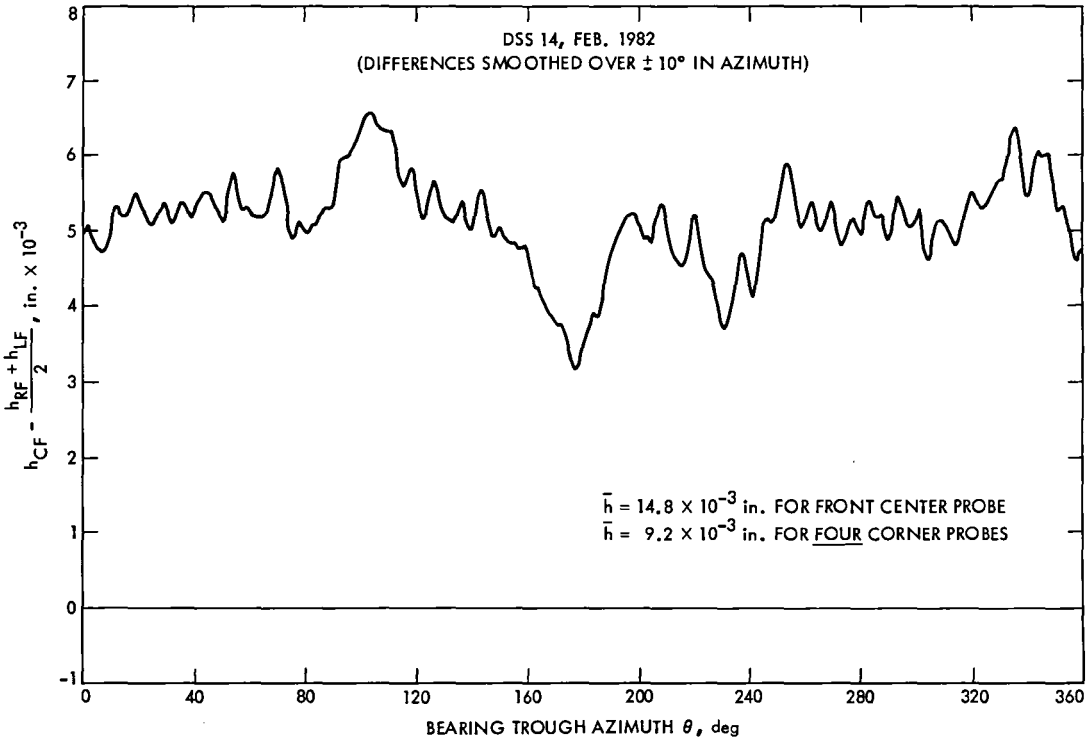


Fig. 12. Film height difference between front center and average of front end probes

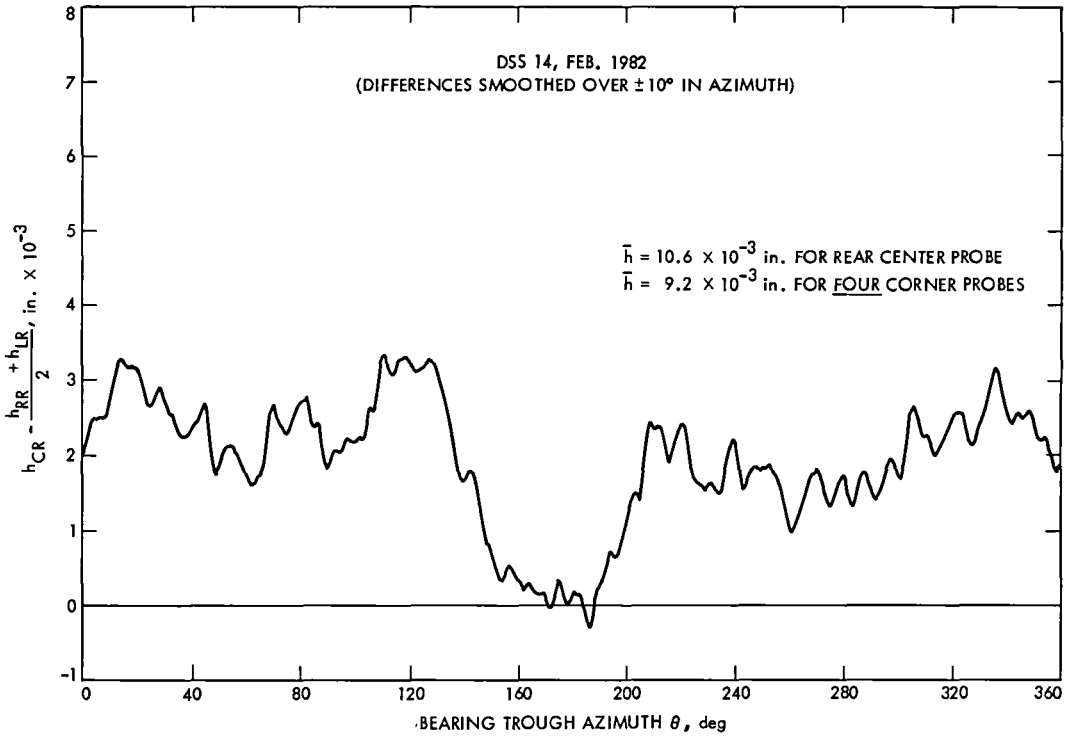


Fig. 13. Film height difference between rear center and average of rear end probes (DSS 14)

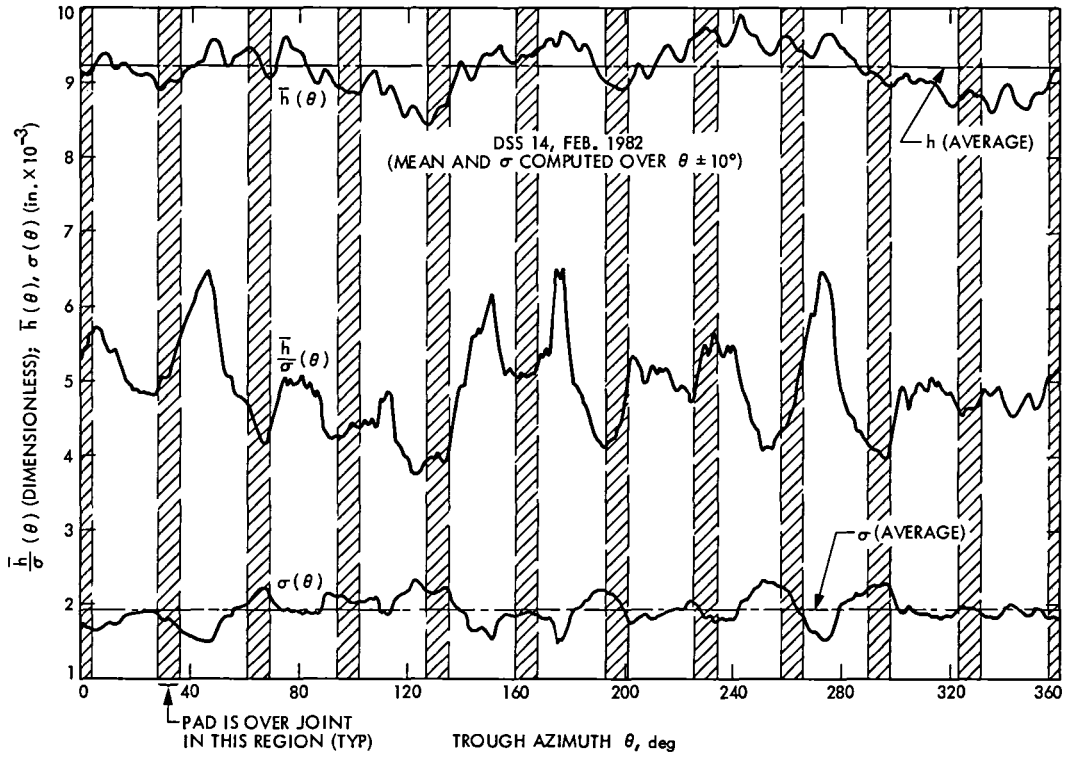


Fig. 14. Plots of local mean film height $\bar{h}(\theta)$, local standard deviation $\sigma(\theta)$, and their quotient $\bar{h}/\sigma(\theta)$ versus trough azimuth (four corner probes)

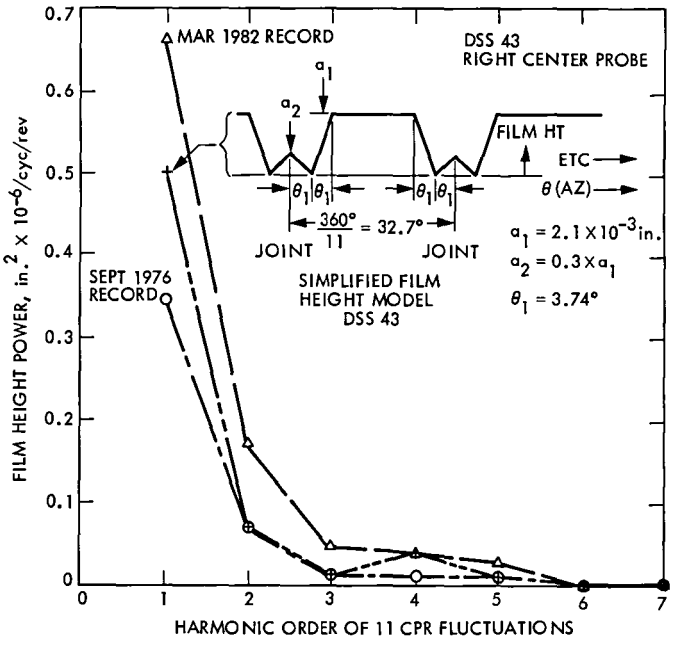


Fig. 15. Power in runner joint induced fluctuations and for simple film height model, DSS 43

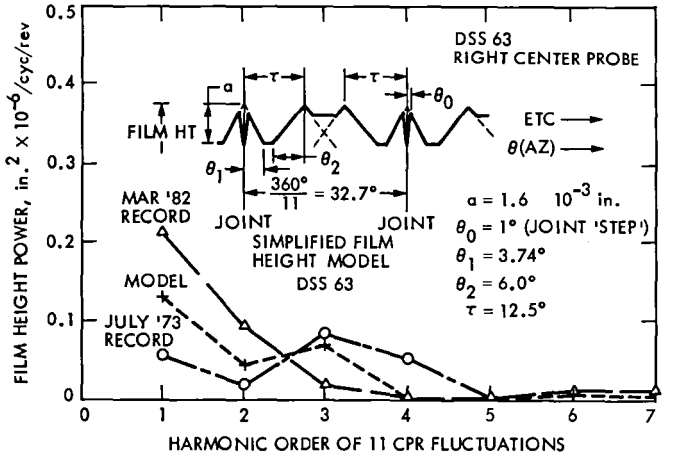
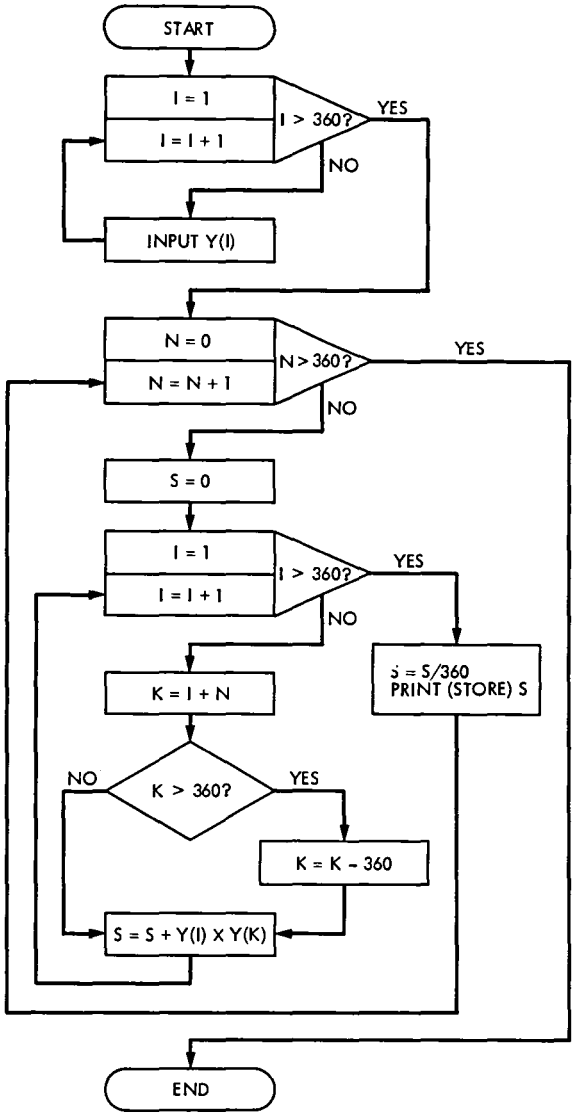


Fig. 16. Power in runner joint induced fluctuations and for simple film height model, DSS 63

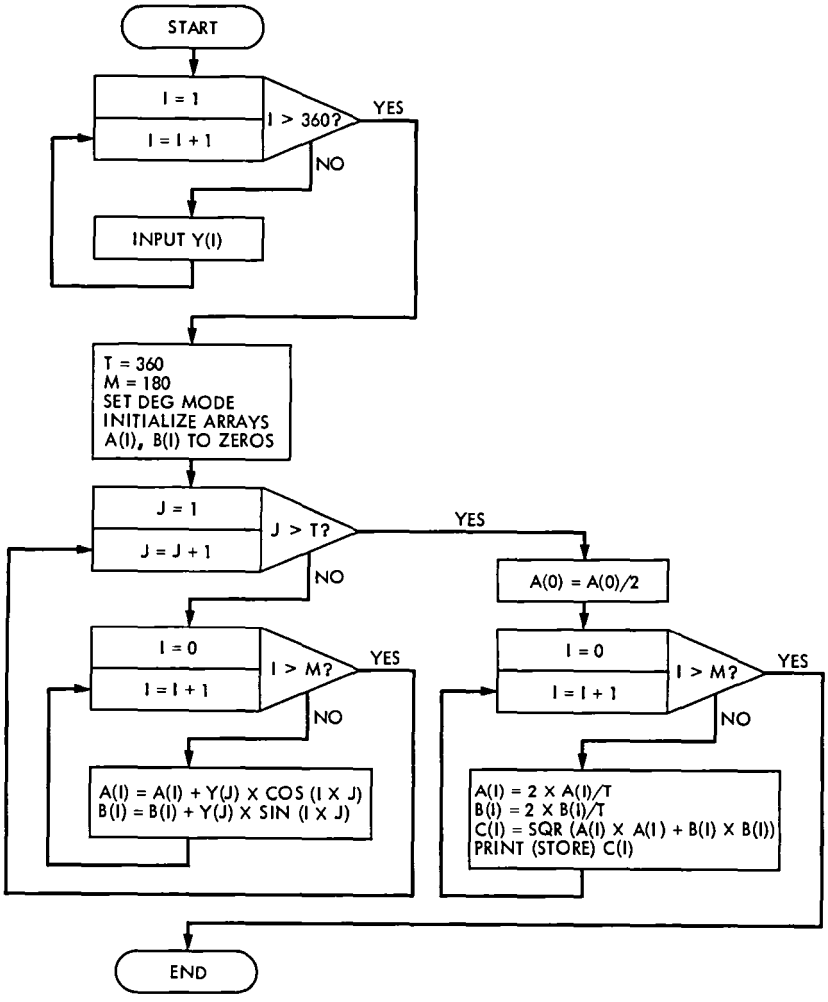
Appendix A

Flow Chart for Calculation of Autocorrelation Function



Appendix B

Flow Chart for Calculation of Power Spectrum from Autocorrelation Function



Appendix C

The Set of Film Height Power Spectra

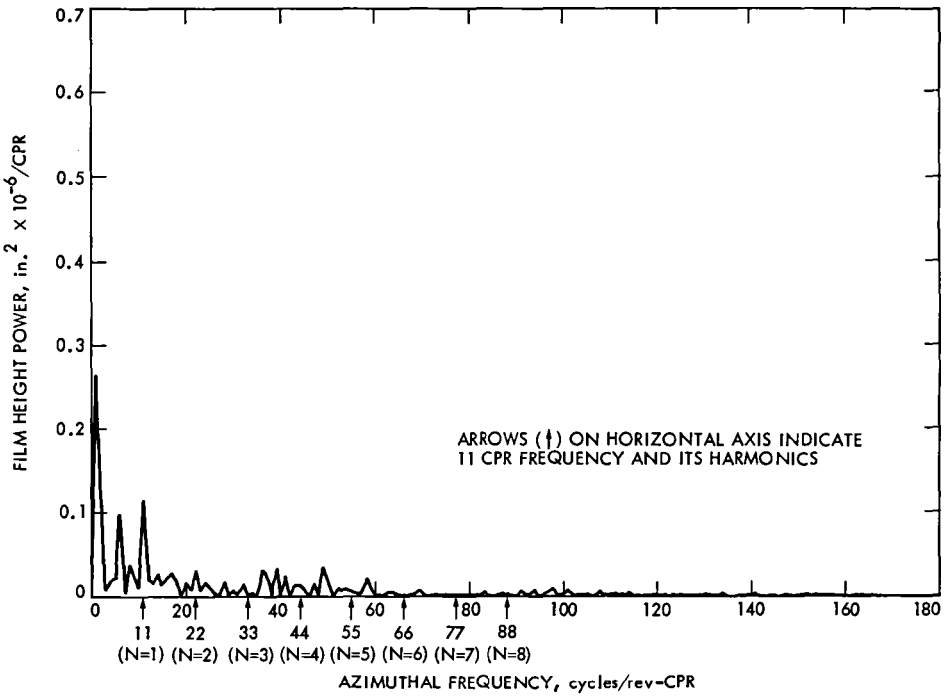


Fig. C-1. Film height power spectrum, right center probe, DSS 14, August 1971

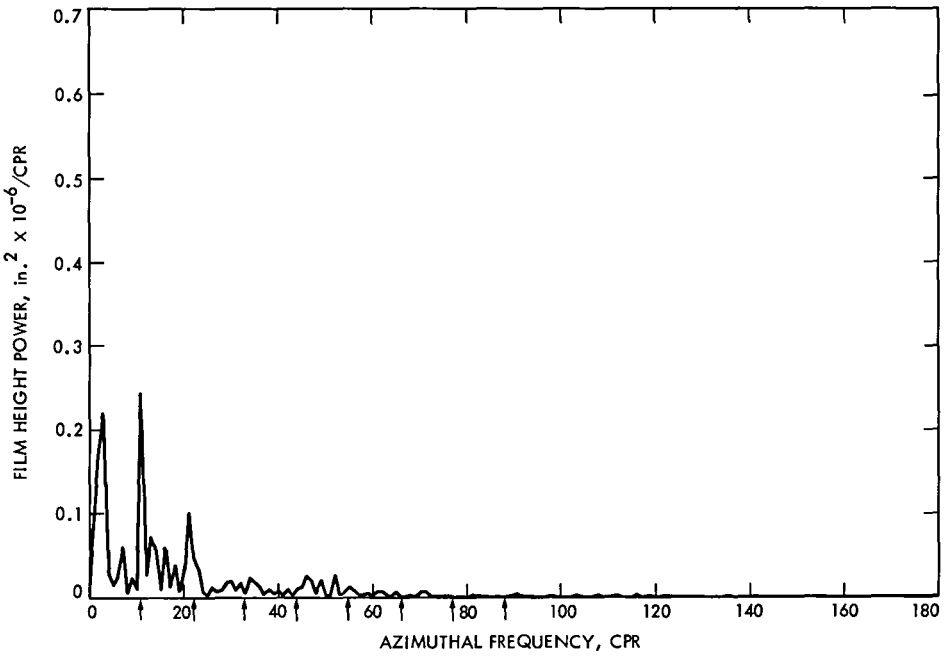


Fig. C-2. Film height power spectrum, right center probe, DSS 14, July 1981

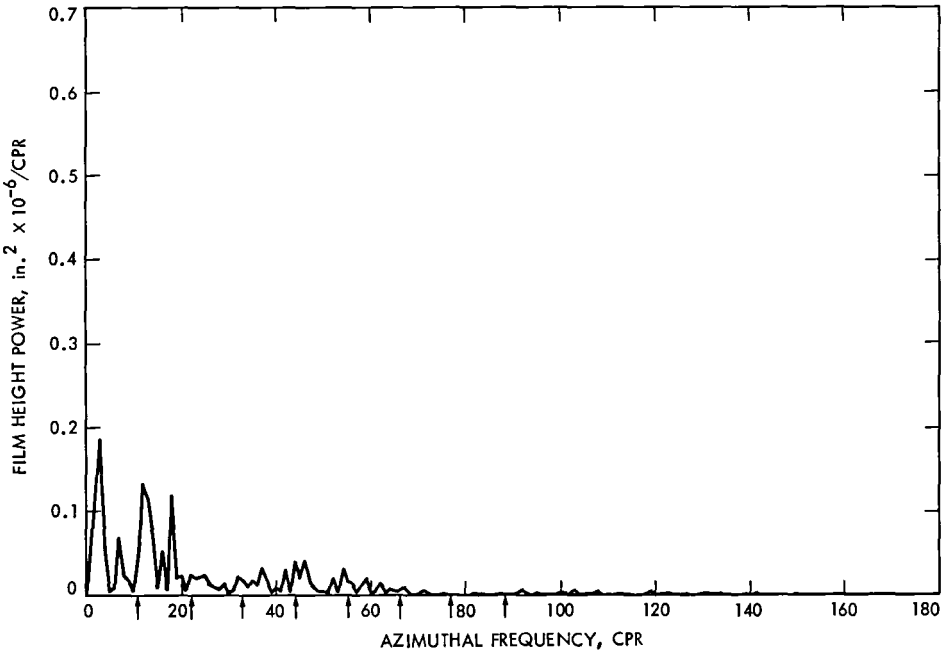


Fig. C-3. Film height power spectrum, right center probe, DSS 14, December 1981

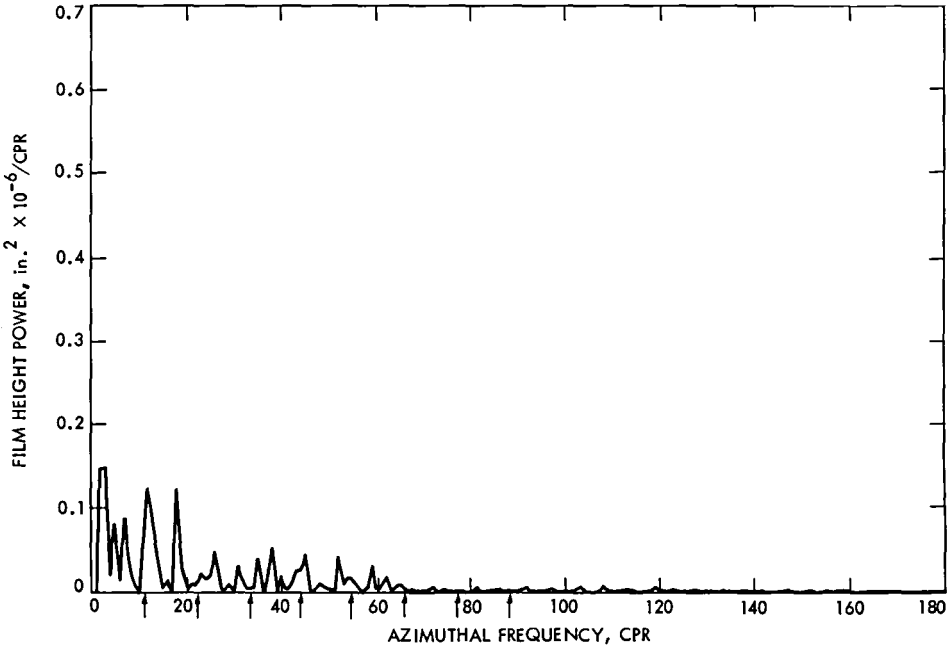


Fig. C-4. Film height power spectrum, right center probe, DSS 14, February 1982

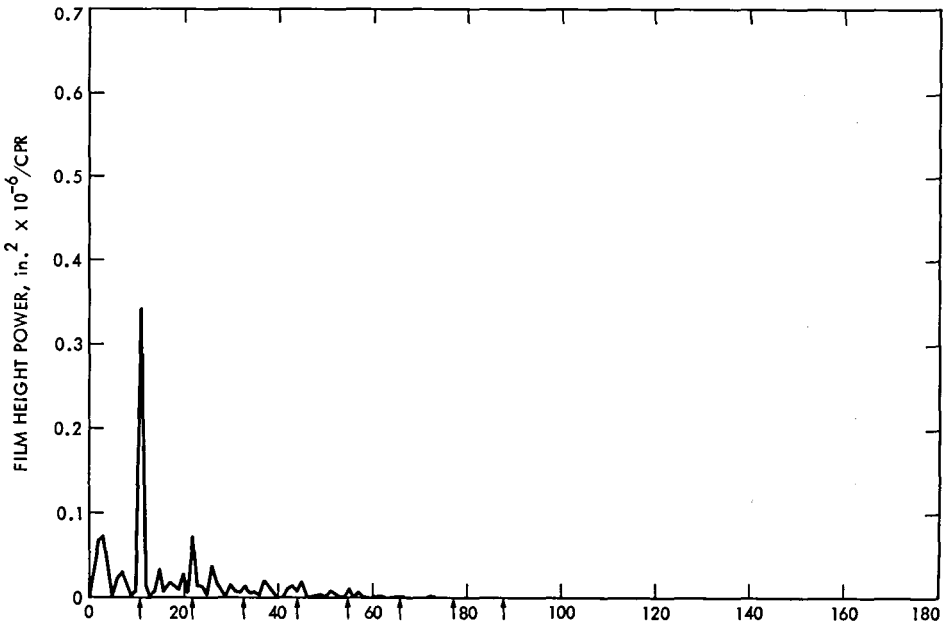


Fig. C-5. Film height power spectrum, right center probe, DSS 43, September 1976

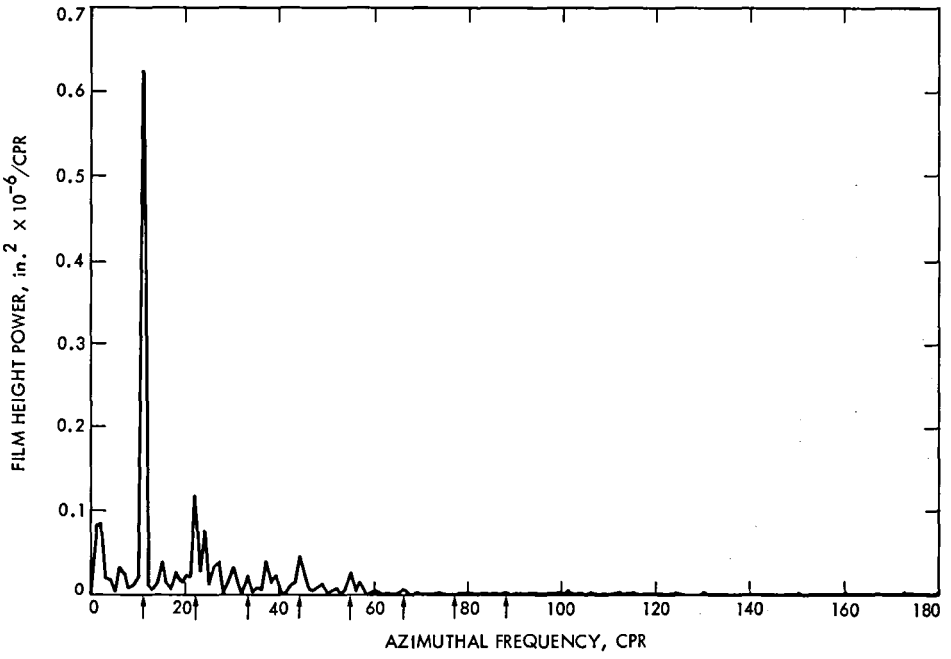


Fig. C-6. Film height power spectrum, right center probe, DSS 43, October 1978

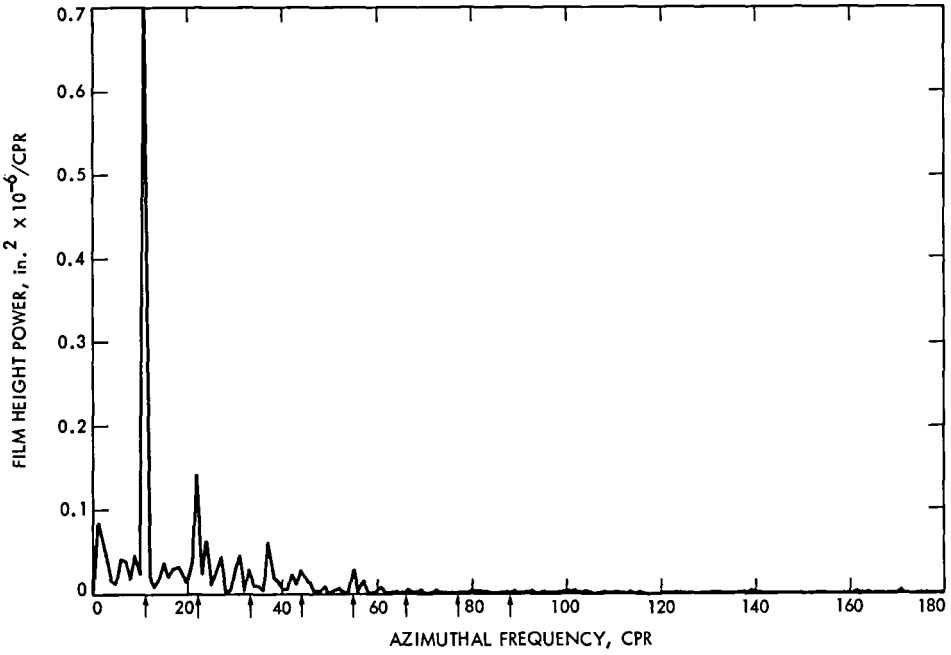


Fig. C-7. Film height power spectrum, right center probe, DSS 43, June 1980

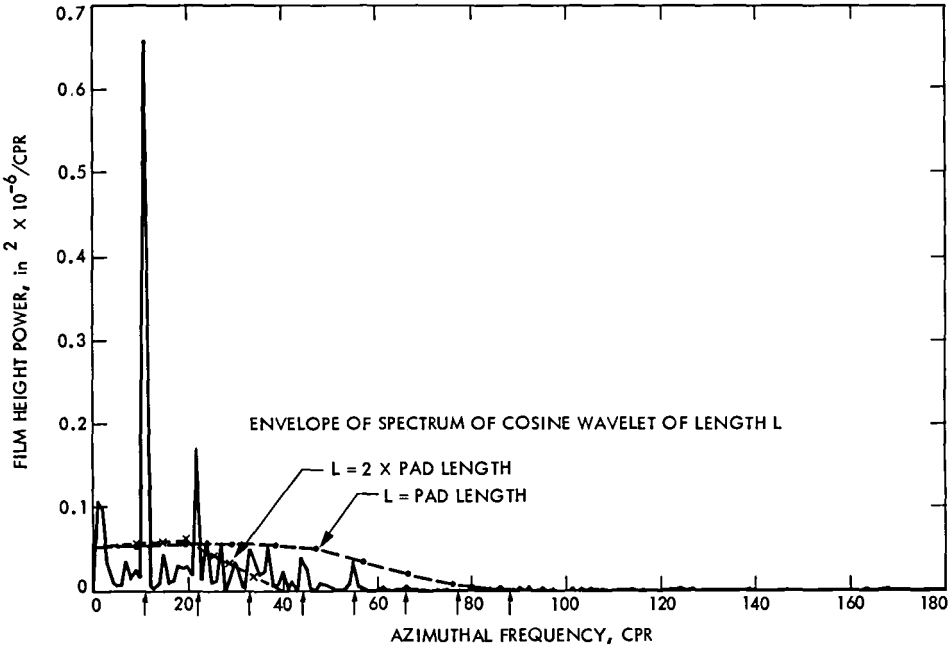


Fig. C-8. Film height power spectrum, right center probe, DSS 43, March 1982

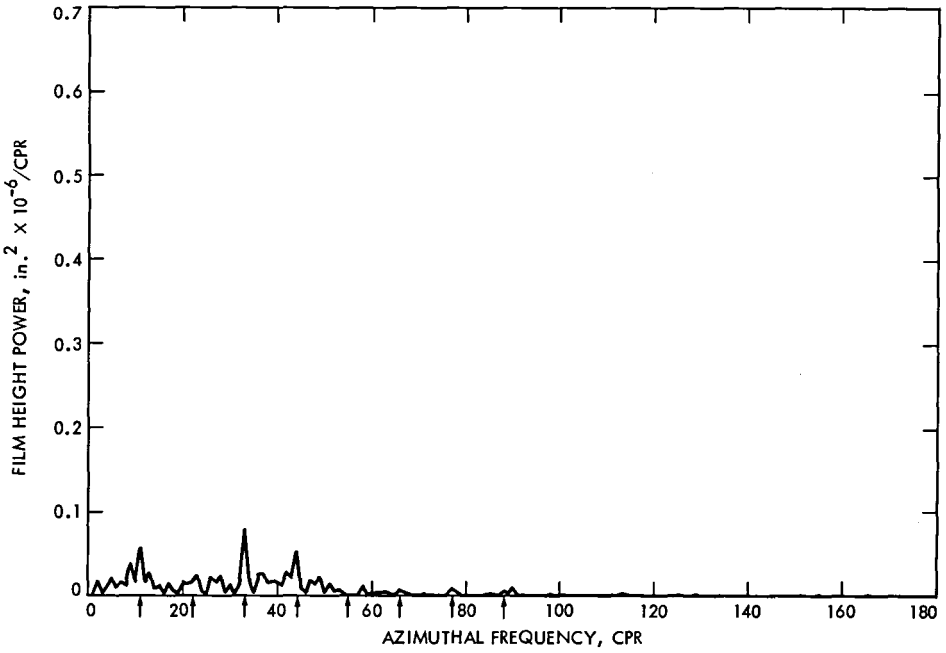


Fig. C-9. Film height power spectrum, right center probe, DSS 63, July 1973

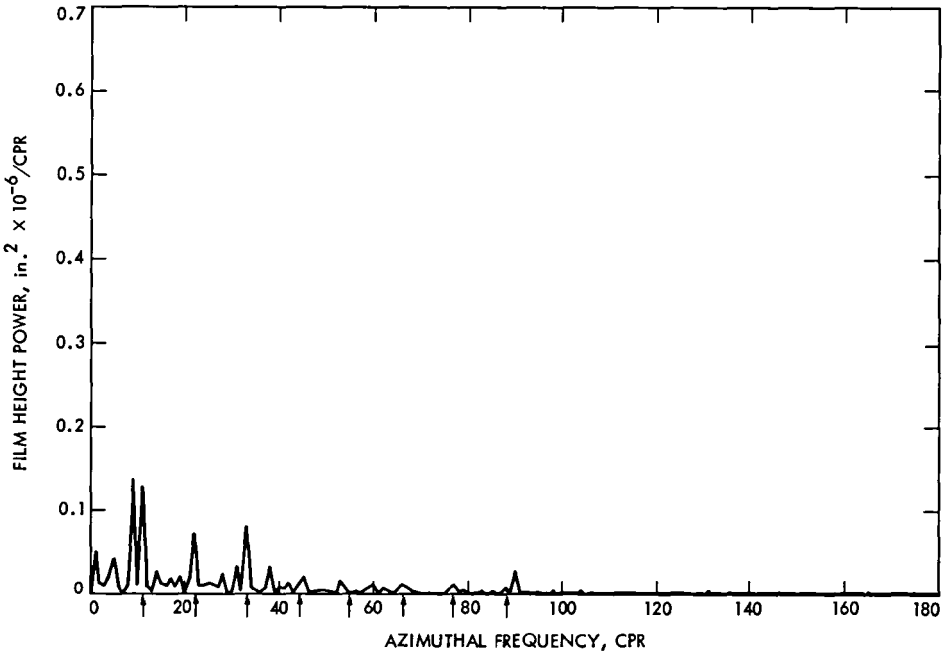


Fig. C-10. Film height power spectrum, right center probe, DSS 63, June 1978

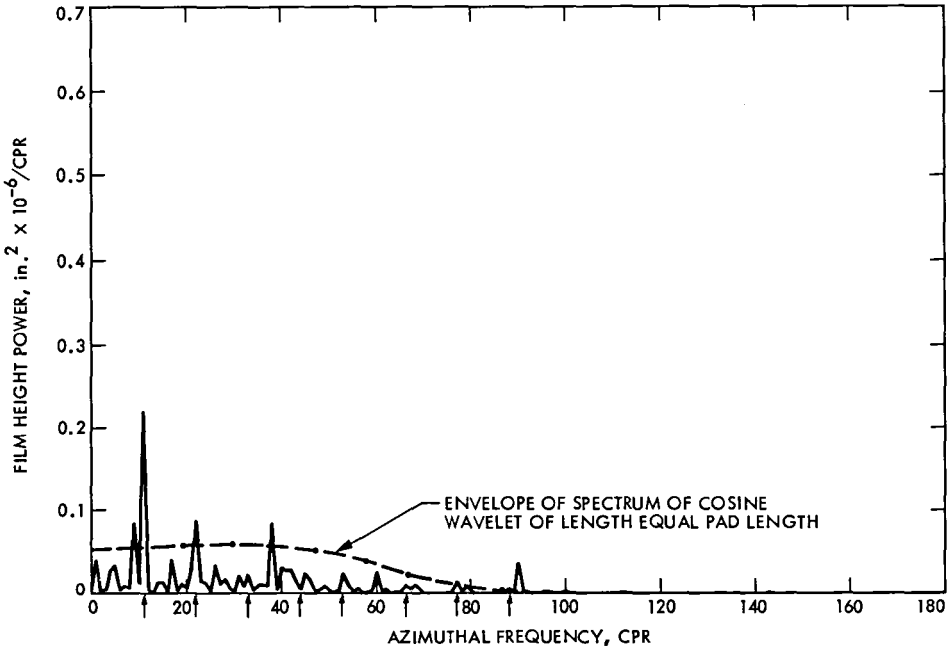


Fig. C-11. Film height power spectrum, right center probe, DSS 63, March 1982

The Application of the Implicit Alternating-Direction Numerical Technique to Thermal Analysis Involving Conduction and Convection

R. D. Hughes and T. Charng
Ground Facilities and Antenna Engineering Section

A computerized model has been developed for analyzing the temperature distribution of a two-dimensional body which is located at or near the soil surface and is partially exposed to solar radiation. The body may have one or more interior cavities containing air or another fluid. The methodology which evolved is also applicable to a general class of thermal analyses involving a body surrounded by a semi-infinite medium exposed to surface radiation energy. The theoretical analysis, numerical procedure, and a sample case are discussed.

I. Introduction

The problem of determining the temperature and heat flux of bodies which are exposed to ambient environmental conditions and have portions lying below the soil surface occurs in a wide variety of applications. Since determining a closed-form analytical solution is difficult and may only be done to a first-order accuracy, the development of a numerical finite-difference model was necessary to determine the temperature distribution of bodies which fall into this class of problems.

The initial study objective was to develop a thermal model of a concrete cable conduit, which is partially built below the surface of the soil, and has an interior cavity. The conduit is subject to solar radiation (insolation) only on its above-ground surface. As the numerical solution was being developed, its usefulness was extended to enable the analysis of a body in a semi-infinite medium having a more general configuration containing one or more cavities, with or without the external surfaces exposed to insolation.

Examples of such general configurations include: a solar pond, an object floating or submerged in a body of water, a microwave antenna pedestal foundation, a covered excavation for cable conduits, an underground tunnel, or a building without exterior glazing.

The heat transfer formulation is presented in Section II, and the numerical procedure in Section III. The finite difference solution employs the implicit alternating-direction method which is an effective compromise between computational accuracy and speed. This method avoids the stringent step-size requirement of the explicit finite difference method and the computational complexity of the two-dimensional implicit method.

II. Thermal Analysis

A. Formulation of Finite Difference Equations

The governing energy balance equation in this two-dimensional thermal analysis is the nonsteady heat conduction equation, which is:

$$c_p \rho \frac{\partial T}{\partial t} = \frac{\partial}{\partial X}\left(k \frac{\partial T}{\partial X}\right) + \frac{\partial}{\partial Y}\left(k \frac{\partial T}{\partial Y}\right) + Q \qquad (1)$$

where

$c_p(X,Y)$ = the local specific heat

$\rho(X,Y)$ = the local density

$T(X,Y,t)$ = temperature

t = time

$k(X,Y)$ = thermal conductivity

$Q(X,Y,t)$ = heat gain (or loss) per unit volume from sources other than conduction, such as internal generation or surface radiation.

In this model, Q consists of surface radiation heat transfer and solar irradiation on the surface.

For a two-dimensional rectangular network of thermal nodes, as shown in Fig. 1, Eq. (1) expressed in the implicit, backward-difference form becomes

$$c_p(i,j)\rho(i,j)\left[\frac{T'(i,j) - T(i,j)}{\Delta t}\right] = k(i,j)\left[\frac{T'(i-1,j) - T'(i,j)}{\Delta X^2} + \frac{T'(i+1,j) - T'(i,j)}{\Delta X^2} + \frac{T'(i,j-1) - T'(i,j)}{\Delta Y^2} + \frac{T'(i,j+1) - T'(i,j)}{\Delta Y^2}\right] + Q(i,j) \qquad (2)$$

at node point (i,j), where the thermal conductivity $k(i,j)$ is considered uniform and constant with temperature within the differential element surrounding (i,j), and Δt is the time step. Note that (i,j) denote position in the X- and Y-directions, respectively, and T' represents the unknown temperatures at time $(t + \Delta t)$.

By analogy to an electrical circuit, the thermal energy balance is thought of in terms of heat flows or "currents" into an element centered about the node point (i,j) with "thermal resistances" connecting all neighboring nodes. Then Eq. (2) becomes:

$$C(i,j)\left[\frac{T'(i,j) - T(i,j)}{\Delta t}\right] = \frac{T'(i-1,j) - T'(i,j)}{R(i-1,j)} + \frac{T'(i+1,j) - T'(i,j)}{R(i+1,j)} + \frac{T'(i,j-1) - T'(i,j)}{R(i,j-1)} + \frac{T'(i,j+1) - T'(i,j)}{R(i,j+1)} + q(i,j) \qquad (3)$$

where

$C(i,j)$ = the thermal capacity of the cell

$q(i,j)$ = the rate of heat gain or loss to the cell from other sources

$R(i,j)$ = typical thermal resistance between node $(i-1,j)$ and node (i,j)

$V(i,j)$ = volume of the cell with unit depth

$\Delta X(i,j)$ = X-dimension of the cell

$\Delta Y(i,j)$ = Y-dimension of the cell

Note that for the two-dimensional case, the depth (Z-dimension) of the cell is taken to be unity, giving

$$V(i,j) = \Delta X(i,j)\,\Delta Y(i,j)$$

$$C(i,j) = V(i,j)\,\rho(i,j)\,C_p(i,j)$$

$$q(i,j) = V(i,j)\,Q(i,j)$$

$$R(i-1,j) = \frac{\Delta X^2\,(i,j)}{k(i-1,j)\,V(i,j)}$$

In dealing with interfaces between different media the following cases are considered for the computation of thermal resistance:

(a) For two adjacent cells having different conductivities, the resistance in the X-direction between node $(i-1,j)$ and node (i,j) is:

$$R(i-1,j) = \frac{1}{\Delta Y(i,j)}\left[\frac{\Delta X(i-1,j)}{2k(i-1,j)} + \frac{\Delta X(i,j)}{2k(i,j)}\right] \tag{4}$$

This relationship applies, for example, to the resistance between nodes (4, 1) and (5, 1) in Fig. 1.

(b) For the case where convection occurs at the boundary between two cells, the resistance between nodes $(i-1,j)$ and (i,j) becomes:

$$R(i-1,j) = \frac{1}{\Delta Y(i,j)}\left[\frac{1}{h(i-1,j)} + \frac{\Delta X(i,j)}{2k(i,j)}\right] \tag{5}$$

where $h(i-1,j)$ is the convective heat transfer coefficient between the $(i-1,j)$ node and the boundary between the cells. Note that Eqs. (4) and (5) are only valid for an interface that is located halfway between the adjacent nodes along the perpendicular direction.

The general form of the thermal resistance equation combining Eqs. (4) and (5) for convection and different material conductivities at thermal interfaces is:

$$R(i-1,j) = \frac{1}{\Delta Y(i,j)}\left[\frac{\Delta X(i-1,j)}{2k(i-1,j)} + \frac{1}{h(i-1,j)} + \frac{\Delta X(i,j)}{2k(i,j)} + \frac{1}{h(i,j)}\right] \tag{6}$$

Appropriate values of thermal conductivity and heat transfer coefficients should be set to match the conditions at each node. "Infinity" values can be used for k or h to selectively reduce appropriate terms to zero in Eq. (6). For instance, if there is convection between the $(i-1,j)$ node and the intercell boundary, then $k(i-1,j)$ should be set to "infinity."

B. Radiation Exchange

The sky-to-surface radiation heat transfer rate per unit area is expressed as:

$$Q_{sky} = \sigma F \epsilon \left[(T_{sky})^4 - T_2^4\right] \tag{7}$$

where

σ = Stefan-Boltzmann constant = 5.6697×10^{-8} W/m^2 K^4

ϵ = $\epsilon(X, Y)$ = effective emissivity of the surface that is exposed to the environment; appropriate values are discussed in Section IV

T_{sky} = the clear sky temperature

T_2 = surface temperature

F = view factor

The sky temperature is assumed as 0.914 T_1, where T_1 = ambient air temperature at the surface.

Equation (7) was applied to flat, horizontal surfaces (where $F = 1$). Other configurations require the determination of radiation shape factor F. The model developed in this report calculates sky-to-surface radiant exchange only for horizontal surfaces.

The rate of solar radiation per unit area absorbed by the surface is represented by the relation:

$$Q_a = \alpha I_t \tag{8}$$

where

Q_a = solar radiation intensity absorbed by a horizontal surface

$\alpha(X, X)$ = average effective solar radiation absorptance of the exposed surface

I_t = incident solar radiation

Thus the net heat gain or loss from other nearby sources, $q(i,j)$ of Eq. (3), combines Eqs. (7) and (8) and is expressed as:

$$q(i,j) = \Delta X(i,j)\left[Q_{sky}(i,j) + Q_a(i,j)\right] \tag{9}$$

for two-dimensional surface cells with unit depth exposed to the external environment. For all interior cells, $q(i,j) = 0$.

III. Numerical Procedure

The finite difference formulation given by Eq. (3) provides a set of N simultaneous linear, algebraic equations for N inte-

rior node points in which the unknowns are values of the updated temperature T'.

In some applications, a special one-dimensional version of Eq. (3) needs to be considered. The equation set in this case can be written in matrix form, and the coefficient matrix becomes tridiagonal. A recursive analytical method is available to solve this tridiagonal system. The following discussion refers to the case which consists of a cable conduit structure buried in soil with the top surface exposed to the environment. More design details are given in the sample case in Section IV.

A. One-Dimensional Analysis

In order to provide a boundary condition for the two-dimensional case, the one-dimensional solution is applied to a vertical section of soil which is sufficiently distant from the cable conduit structure so that conduction in the horizontal direction is negligible. The resulting temperature profiles may then be taken as time-dependent boundary conditions for the appropriate boundary of the two-dimensional case.

The one-dimensional form of Eq. (3) is:

$$C(j)\frac{T'(j)-T(j)}{\Delta t} = \frac{T'(j-1)-T'(j)}{R(j-1)} + \frac{T'(j+1)-T'(j)}{R(j+1)} + q(j)$$

$$(q(j) = 0 \quad \text{for } j > 1) \tag{10}$$

which may be expressed as:

$$-\frac{1}{R(j-1)}T'(j-1) + \left[\frac{C(j)}{\Delta t} + \frac{1}{R(j-1)} + \frac{1}{R(j+1)}\right]T'(j)$$

$$-\frac{1}{R(j+1)}T'(j+1) = \frac{C(j)T(j)}{\Delta t} + q(j) \tag{11}$$

The coefficients of $T'(j-1)$, $T'(j)$, and $T'(j+1)$ designates $a(j)$, $b(j)$, and $c(j)$ respectively form a tridiagonal matrix with diagonal vectors a, b, and c, where

$$a(j) = -\frac{1}{R(j-1)}$$

$$b(j) = \frac{C(j)}{\Delta t} + \frac{1}{R(j-1)} + \frac{1}{R(j+1)}$$

$$c(j) = \frac{1}{R(j+1)}$$

The right-hand side of Eq. (11) is a known quantity and is designated as the vector d where:

$$d(j) = \frac{C(j)T(j)}{\Delta t} + q(j)$$

The recursion solution (Ref. 2) is of the form:

$$T'(j) = \gamma(j) - \frac{c(j)}{\beta(j)}T'(j+1) \qquad j = N-1, N-2, \cdots, 1$$

$$T'(N) = \gamma(N) \tag{12}$$

where

$$\beta(j) = b(j) - \frac{a(j)c(j-1)}{\beta(j-1)} \qquad j = 2, 3, \cdots, N \tag{13}$$

$$\gamma(j) = \frac{d(j) - a(j)\gamma(j-1)}{\beta(j)} \qquad j = 2, 3, \cdots, N \tag{14}$$

$$\beta(1) = b(1) \quad \text{(node 1 is the surface node)} \tag{15}$$

$$\gamma(1) = \frac{d(1)}{\beta(1)} \tag{16}$$

By this approach, the problems of accumulation of round-off error or iterative convergence which are typical with other methods of matrix equation solution are eliminated.

In applying Eqs. (12-16), the initial and boundary conditions may be introduced by specifying an initial temperature at each node, at time ($t = 0$) and boundary temperatures known as a function of time at nodes $j = 0$ and $j = N + 1$, which represent the ambient air and a "constant" deep earth location, respectively. For this analysis, the ambient air temperature, $T(i,0)$, as a function of time of day was approximated as a sinusoidal function having a given average, amplitude, and phase. These given quantitites were determined by curve-fitting to local recorded temperature data for the site being modelled. The one-dimensional analysis was taken to a depth such that the temperature at the bottom node, $j = N + 1$, was essentially constant on both a diurnal and annual basis. Surface effects typically penetrate soil to a depth of about 20 meters throughout the time span of the annual cycle. A separate computer program has been written to perform this one-dimensional analysis.

Since modelling soil temperatures is a cyclic problem, the one-dimensional model has to be run for each day of the year in succession, for about five annual cycles to remove the effects of initial temperature conditions.

B. Two-Dimensional Numerical Procedure

When the two-dimensional form of Eq. (3) is considered, the coefficient matrix is found to no longer be tridiagonal,

and the recursive solution cannot be used to solve the system of equations. However, the implicit alternating-direction method again allows use of the tridiagonal recursion solution and avoids the difficulty of solving the matrix equation by other methods. The implicit alternating-direction method uses two different equations in turn over successive time-steps, each of duration $\Delta t/2$. The first equation is implicit only in the X-direction between times t and $(t + \Delta t/2)$ and the second equation in the Y-direction between times $(t + \Delta t/2)$ and $(t + \Delta t)$ as follows:

$$-\frac{1}{R(i-1,j)}T^*(i-1,j)+\left[\frac{C(i,j)}{\Delta t/2}+\frac{1}{R(i-1,j)}+\frac{1}{R(i+1,j)}\right]T^*(i,j)-\frac{1}{R(i+1,j)}T^*(i+1,j)=\frac{1}{R(i,j-1)}T(i,j-1)$$

$$+\left[\frac{C(i,j)}{\Delta t/2}-\frac{1}{R(i,j-1)}-\frac{1}{R(i,j+1)}\right]T(i,j)+\frac{1}{R(i,j-1)}T(i,j-1)+q(i) \quad (17)$$

and

$$-\frac{1}{R(i,j-1)}T'(i,j-1)+\left[\frac{C(i,j)}{\Delta t/2}+\frac{1}{R(i,j-1)}+\frac{1}{R(i,j+1)}\right]T'(i,j)-\frac{1}{R(i,j+1)}T'(i,j+1)=\frac{1}{R(i-1,j)}T^*(i-1,j)$$

$$+\left[\frac{C(i,j)}{\Delta t/2}-\frac{1}{R(i-1,j)}-\frac{1}{R(i-1,j)}\right]T^*(i,j)+\frac{1}{R(i+1,j)}T^*(i+1,j)+q(i) \quad (18)$$

where T is temperature at time t, T^* is temperature at time $(t + \Delta t/2)$, T' is temperature at time $(t + \Delta t)$. The right-hand sides of Eqs. (17) and (18) are known quantities and the temperatures on the left-hand sides are the unknowns. By using tridiagonal matrix techniques, Eq. (17) is solved for the intermediate values T^*, which are then used in Eq. (18) to similarly solve for T' at the end of the time interval Δt.

The boundary conditions are prescribed along each of the four boundaries as follows:

(1) The top boundary condition is the temperature of ambient air above the soil surface as a function of time. The air temperature is determined as previously described for the one-dimensional case.

(2) The right-hand boundary is the temperature profile of homogenous soil as a function of time. The results of the one-dimensional analysis are used for this boundary condition.

(3) The left-hand boundary is the vertical line of symmetry of the structure. A symmetric boundary condition implies thermal gradients of zero in the normal direction to the boundary.

(4) The bottom boundary condition is a prescribed temperature which is constant along the bottom row of nodes, but can vary with time. The value is taken from the corresponding node on the right-hand boundary at each time step. This boundary must be located at a sufficient depth so that it can be assumed unaffected by the thermal influence of the structure. A trial-and-error approach may be necessary here to avoid unacceptably large nodal networks.

Although the boundary conditions have been included in the computer model in the form described, they can be modified to fit another problem with little difficulty.

IV. Sample Case

The sample case configuration consisted of a hypothetical concrete trench designed to house communication cables which require a thermally stable environment. The trench, shown in cross-section in Fig. 2, is covered by steel plates which are coated with a highly reflective paint and lined with styrofoam on the inside. Ambient conditions at the Goldstone Deep Space Communications Complex were used for this case.

A. One-Dimensional Analysis

The incident solar radiation I_t is computed by a subroutine, SOLAR, which uses the ASHRAE formulation described in Ref. 1 to compute direct and diffuse sky components. Input

parameters to this subroutine include site latitude, day of the year, and angle of the surface with respect to the horizon (zero in this model). Diffuse sky radiation and direct solar radiation are the only components which are included in computing I_t. Solar radiation reflected from nearby surroundings can be determined by the SOLAR subroutine, but is not used in this model.

The input conditions for the one-dimensional analysis were: latitude = 34°N, effective surface absorptivity $\alpha = 0.4$, surface emissivity $\epsilon = 0.45$, time step = 1 hour, initial temperature = 20.6°C, soil density $\rho = 2.05$ g/cm^3, soil specific heat = 1.84 J/g-k, soil conductivity k = 9.5×10^{-3} W/cm-K, surface convection coefficient h = 2.27×10^{-3} W/cm^2-K; for ambient air temperature: annual average = 18.9°C, amplitude of annual temperature wave = 11.4°C, daily amplitude = 12.2°C, and the constant soil temperature at a depth of 17 m was taken to be 20.6°C.

No cloud cover factors were used, and no direct means was included to take into account the precipitation-evaporation cycle which causes heat loss at the soil surface. Thus an "effective" absorptivity was defined which accounts for these factors. A typical set of calculated temperature profiles is shown in Fig. 3, indicating conditions at four different times on a particular day. The left-most point on the curves is the ambient air temperature. Note that the effect of diurnal fluctuation in ambient temperature is damped to nearly zero at a depth of about 0.5 m below the surface. The remainder of the profile is characterized by annual temperature cycle effects. The validity of the model result is confirmed by the fact that the average temperature profile for the entire year was effectively uniform and equal to the constant soil temperature of 20.6°C, and the behavior of the solution in terms of time lag and damping of temperature amplitude as a function of depth is analogous to published analytical solutions.

B. Two-Dimensional Analysis

The additional input conditions for the two-dimensional analysis consisted of handbook values for thermal properties of the trench materials, absorptivity of the cover plate = 0.1, and emissivity of the cover plate = 0.12. The analysis was done for the 170th day of the year. The model was cycled for three 24-hour periods to overcome the effect of initial conditions (all the columns of nodes were initially set equal to the right-hand-side boundary condition). Figure 4 shows the nodal network configuration which was input to the computer model.

Results are shown at times 0100 and 1200 PST in Table 1. Ten nodes were used in the vertical direction and eight in the horizontal. Note that the temperatures in row 9 are nearly uniform, showing that the assumption of uniform temperature along the bottom boundary (row 10) is valid. Also, temperatures in the horizontal direction in columns 5-8 are nearly uniform, showing that the assumption of one-dimensional heat transfer for the right-hand boundary is valid.

V. Discussion

The modelling procedure outlined in this report is applicable to a general class of problems and offers several advantages over using large general-purpose thermal analysis computer programs. Among these advantages are the compact size of the software, the economy of operation of the software, and the simplicity of applying the modelling procedure. However, there are some restrictions inherent in the model as it presently exists: (1) the thermal node mesh must be rectangular, (2) the left-hand boundary must be a line of symmetry, (3) radiant heat exchange inside the cavity is neglected, and (4) no cloud cover factors are included in the solar model.

All of these restrictions may be overcome without great difficulty by additional development of the modelling software. This thermal model will become an increasingly general analytical tool when these additional capabilities are implemented.

References

1. Hughes, R. D., "A Simplified Solar Cell Array Modelling Program," *TDA Progress Report 42-68*, pp. 167-185, Jet Propulsion Laboratory, Pasadena, Calif., Apr. 15, 1982.
2. Carnahan, B., Luther, H. A., and Wilkes, J. O., *Applied Numerical Methods*, John Wiley & Sons, Inc., N.Y., 1969.

Table 1. Soil temperatures for Day 170 at two representative times (Two-dimensional temperature profiles, °C)

Time = 0100								
Vertical node	Horizontal node							
	1	2	3	4	5	6	7	8
1	25.63	25.63	25.63	25.63	25.63	25.63	25.63	25.63
2	27.03	24.42	27.03	26.71	26.22	26.30	26.30	26.88
3	30.60	26.33	30.60	31.57	31.71	31.78	31.77	31.61
4	30.07	27.04	30.07	29.76	29.91	29.98	30.00	29.83
5	27.25	26.40	27.25	27.20	27.28	27.31	27.30	27.13
6	25.09	24.10	25.09	25.08	25.09	25.09	25.07	24.88
7	23.51	22.54	23.51	23.57	23.59	23.59	23.57	23.40
8	22.79	22.28	22.79	22.81	22.82	22.82	22.80	22.65
9	21.87	21.75	21.87	21.88	21.88	21.88	21.86	21.70
10	20.74	20.74	20.74	20.74	20.74	20.74	20.74	20.74
Time = 1200								
Vertical node	Horizontal node							
	1	2	3	4	5	6	7	8
1	28.80	28.80	28.80	28.80	28.80	28.80	28.80	28.80
2	30.55	32.79	30.55	37.20	37.48	37.42	37.43	36.79
3	27.28	30.44	27.28	28.90	29.41	29.42	29.44	29.59
4	27.51	28.10	27.51	28.54	28.91	28.98	28.98	28.92
5	27.12	25.67	27.12	27.24	27.31	27.35	27.33	27.17
6	25.12	23.58	25.12	25.12	25.13	25.13	25.10	24.93
7	23.52	22.45	23.52	23.60	23.63	23.63	23.60	23.44
8	22.81	22.22	22.81	22.84	22.86	22.86	22.83	22.69
9	21.89	21.73	21.89	21.90	21.91	21.92	21.88	21.74
10	20.78	20.78	20.78	20.78	20.78	20.78	20.78	20.78

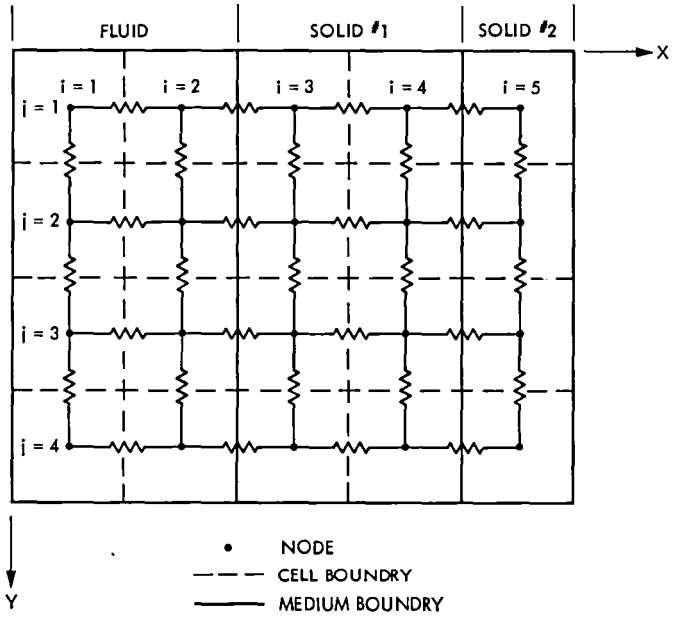
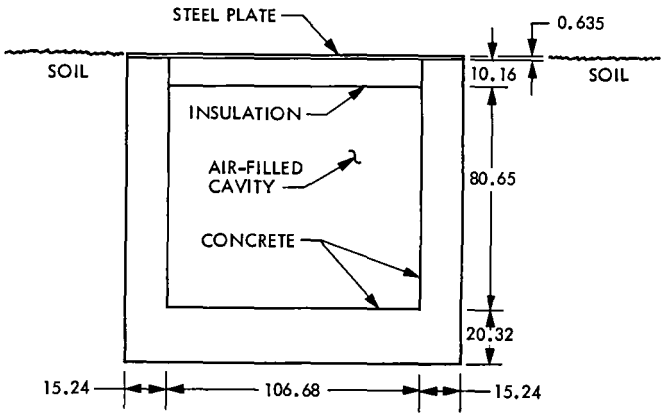


Fig. 1. Nodal representation of two-dimensional heat conduction



NOTE: ALL DIMENSIONS ARE IN CENTIMETERS

Fig. 2. Cable trench configuration

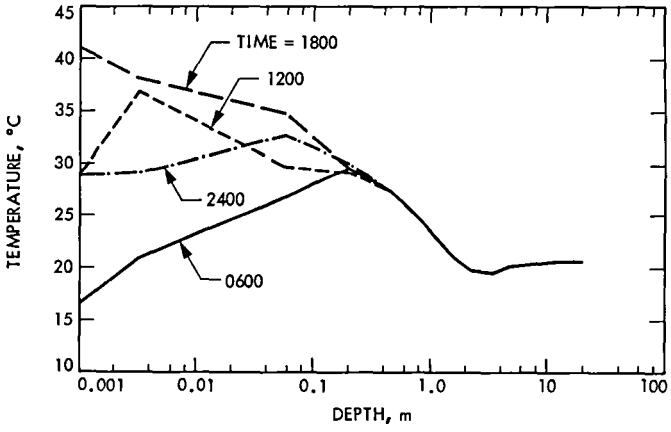
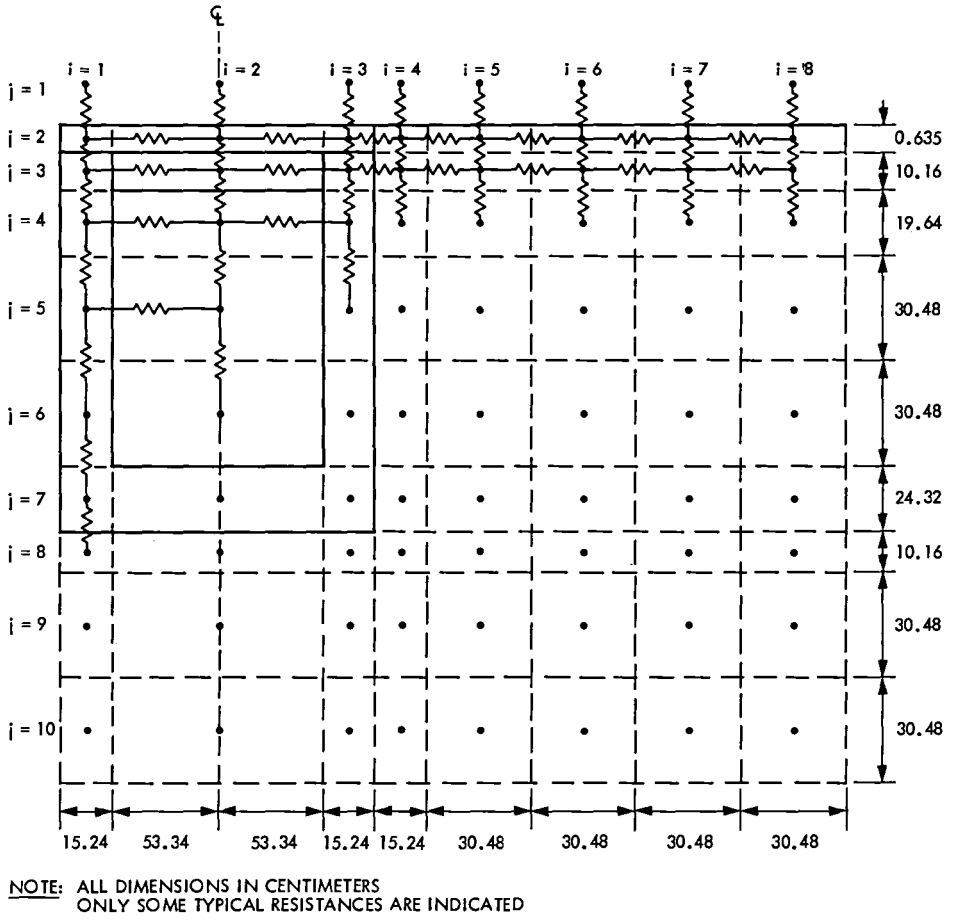


Fig. 3. One-dimensional soil temperature profiles, Day 170



NOTE: ALL DIMENSIONS IN CENTIMETERS
ONLY SOME TYPICAL RESISTANCES ARE INDICATED

Fig. 4. Nodal representation of cable trench and surrounding soil

Radio Interferometric Determination of Source Positions, Intercontinental Baselines, and Earth Orientation With Deep Space Network Antennas – 1971 to 1980

J. B. Thomas, O. J. Sovers, J. L. Fanselow,
E. J. Cohen, G. H. Purcell, Jr., D. H. Rogstad,
L. J. Skjerve and D. J. Spitzmesser
Tracking Systems and Applications Section

A series of experiments has been conducted at the Jet Propulsion Laboratory during the last decade to develop a radio interferometric system capable of measuring crustal and rotational motions of the earth, as well as source positions for a reference frame based on compact extragalactic radio sources. With the exception of one session between Big Pine, Calif., and Westford, Mass., the observing stations were those of NASA's Deep Space Network in California, Spain, and Australia.

Approximately 2400 observations of extragalactic radio sources were made between August 1971 and February 1980 during 48 separate sessions. These consisted of 259 delay rate observations at 2.3 GHz (S-band), 796 delay and delay rate observations at either S-band or 8.3 GHz (X-band) and 1325 delay and delay rate observations recorded simultaneously at both S- and X-band. A single multiparameter fit has been applied to the observed values of delay and delay rate to extract astrometric and geophysical parameters from this decade-long sequence. The fit produced estimates of 784 parameters, including station locations, radio source positions, polar motion, Universal Time, the precession constant, and solid earth tides. The a priori model included gravitational bending, the 1980 IAU nutation series, the 1976 IAU expressions for Greenwich mean sidereal time and precession, BIH estimates of Universal Time and polar motion, and monthly mean values for zenith troposphere delay.

The rms residuals were 0.52 nsec for delay and 0.30 psec/sec for delay rate. Intercontinental baseline lengths have been determined with formal uncertainties of 5 to 10 cm. Universal Time and polar motion were measured at 49 epochs, with formal uncertainties (for the more recent data) of 0.5 msec for UT1 and 6 and 2 mas, respectively, for the X and Y components of polar motion. Our 1971–80 data produced a formal estimate of the luni-solar precession constant that was smaller than the 1976 IAU value by −3.8 ±

0.9 mas/year. However, due to the relatively short span of data and the inaccuracy of the earlier data, the precession effect could not be separated from the effect of the 18.6-year nutation term. Further, the less reliable data in the earlier years may have biased our precession result by an amount considerably larger than the formal error of 0.9 mas/year. The result for the gamma factor of the parameterized post-Newtonian formalism was 0.997 ± 0.041, which agrees very well with general relativity. The vertical and horizontal Love numbers have been determined with 5% and 30% uncertainties, respectively, while the earth tide phase lag was found to be zero within its error estimate. These earth-tide results agree with the commonly accepted values. In addition to these geophysical results, the positions of 104 sources have been obtained with formal uncertainties of approximately 5 milliarcseconds.

I. Introduction

Over the last few years, considerable progress (Refs. 1–5) has been made toward realizing the potential of radio interferometry for measuring local crustal and global rotational motions of the earth with accuracies at the centimeter level. Toward this goal, a series of experiments, primarily with NASA's Deep Space Network (DSN) antennas, has been conducted over the last decade to develop two generations of very long baseline interferometric (VLBI) systems. In all, 48 interferometric sessions were carried out between eight different antennas on three continents. Delay and/or delay rate observables were measured on two local baselines (at Goldstone, California, and at Madrid, Spain), on a transcontinental baseline (Big Pine, California to Westford, Massachusetts) and two intercontinental baselines (Goldstone, California to Madrid, Spain, and Goldstone, California to Tidbinbilla, Australia). A single multiparameter fit has been applied to this decade-long sequence of observations to extract significant astrometric and geophysical parameters. The adjusted parameters included station locations, source positions, polar motion, Universal Time, the precession constant, solid earth tides, and the gamma factor of relativity theory. This report outlines the techniques, analyses and results of these experiments.

II. Interferometry Technique

In interferometry measurements, the random broadband emission of an extragalactic radio source is simultaneously recorded at two widely separated radio antennas. Cross correlation of the recorded data at a central site leads to a determination of the difference in arrival times of the radio wavefront at the two antennas. Since this difference (delay) depends on the direction of the source and the vector separation of the antennas, it contains information concerning the rotational and crustal motions of the earth. To extract this information, it is necessary to measure the delay for many different sources, preferably on several baselines, and to pass the resulting delays through a multiparameter fitting program that solves for astrometric and geophysical quantities.

In the present experiments, two separate interferometry systems have been used to measure delay and delay rate. Since the instrumentation and data processing techniques have been described elsewhere (Refs. 6, 7), we will only summarize the most salient features of the two systems. The prototype system used in the early measurements recorded a single, narrowband (24-kHz) channel at S-band (2.3 GHz) and therefore could measure only delay rate accurately. Since the delay rate observable is independent of the polar component of the baseline vector and is less powerful in parameter estimation than delay, another system was developed to measure both delay and delay rate. This system records several time-multiplexed frequency channels so that delay can be obtained by means of bandwidth synthesis, a technique pioneered by Rogers (Ref. 8). The bandwidth of the individual channels was also increased to 2 MHz to improve flux sensitivity and signal-to-noise ratio. Bandwidth limitations at the first stage of amplification in the Deep Space Network receiver systems limited the maximum bandwidth spanned by the outer channels in a given radio frequency (RF) band to about 40 MHz. With this bandwidth synthesis system, delay could be measured with a precision (i.e., system noise error) of approximately 100 psec, given a typical source strength of 0.5 Jy, an integration time of 3 min, and two 64-m DSN antennas with system temperatures of 35 K. By cycling through two sets of channels properly placed at RF, the second system was capable of "simultaneously" measuring delay at S-band and X-band (8.4 GHz).

III. Geometric Delay Model

The delay (or delay rate) observable measured by radio interferometry is a sum of the differential delays (or rates) due to geometric, instrumental, ionospheric, and tropospheric effects. This section summarizes a mathematical model for the geometric delay that was developed by Fanselow (Ref. 9).

The geometric delay, which is by far the largest (up to 30 msec) of the delay terms, is of primary interest in the current measurements since it depends on the direction of the incom-

ing wave and on the time-varying baseline vector between antennas. Consequently, a calculation of the geometric delay must include, in addition to the radio source location, all significant factors describing the rotational, crustal, and orbital motions of the earth.

Models of the geometric delay usually place the origin of coordinates at either of two points: the center of mass of the earth (geocentric approach) or the center of mass of the solar system (solar-system-barycentric approach). Even though the approximate delay model generated by the geocentric approach is mathematically simpler in its final form, the solar-system-barycentric (SSB) approach is inherently more adaptable to refinements in the geometric delay model, particularly for relativistic effects. Consequently, even though such refinements are not justified by the limited accuracy of the present data, the SSB approach has been adopted in anticipation of model improvements that will be demanded by future, more accurate data. In other applications, the SSB approach would facilitate generalization to accommodate sources (such as spacecraft) within the solar system.

A geometric delay model that adequately describes the present data can be based on the following theoretical assumptions. Given a set of SSB coordinates defined in terms of the mean equator of J2000.0, suppose that a wavefront with propagation direction $\hat{\mathbf{k}}$ is received by two earth-fixed antennas whose time-varying positions in SSB coordinates are given by $\mathbf{x}_1(t)$ and $\mathbf{x}_2(t)$ at TDB (barycentric dynamical) time t. In terms of these quantities, the geometric delay between antennas will be given approximately by

$$\tau_g(t) = \frac{\hat{\mathbf{k}} \cdot [\mathbf{x}_2(t) - \mathbf{x}_1(t)]}{1 + \dfrac{\hat{\mathbf{k}} \cdot \mathbf{v}_2(t)}{c}} \qquad (1)$$

where the dot product involving $\mathbf{v}_2$ (the velocity of antenna 2) accounts for motion of antenna 2 during the wave transit. Since the measured delay is obtained by earth-fixed observers, the theoretical delay in Eq. (1) must be relativistically transformed to earth-fixed coordinates. For the present data, this transformation is adequately applied with a special relativistic (Lorentz) transformation that introduces, as its largest effect, the equivalent of the well-known "earth-centered" aberration correction to the source location. General relativistic effects, except for gravitational bending of radio waves, are neglected. A correction term is applied to the SSB model delay to account for such gravitational bending prior to transforming back to an earth-centered coordinate system.

To obtain the SSB antenna "trajectories" $\mathbf{x}_i(t)$ used in the above delay calculation, a model for the rotational and orbital motions of the earth must be adopted. Our present rotational model for the earth is the model summarized by Kaplan (Ref. 10), including the earth's spin axis direction from the IAU 1976 precession (Ref. 11) and the IAU 1980 expressions for nutation (Refs. 12, 13) and Greenwich mean sidereal time (Ref. 14). Orientation about the spin axis, parameterized as Universal Time (UT1), is treated as a solve-for parameter, as is polar motion (the position of the ephemeris pole with respect to the earth's crust). Station locations are defined relative to an earth-fixed frame with its Z axis along the mean spin axis of 1903.0 and X axis along the Greenwich meridian. Within this frame, generally referred to as the CIO (Conventional International Origin) frame, the station locations are expressed in terms of cylindrical coordinates. Axis orientation for the earth-fixed CIO frame is experimentally determined through the use of Universal Time and polar motion (UT/PM) values measured by the Bureau International de l'Heure (BIH), as explained in Sec. VII. Since all these conventional earth orientation parameters are defined relative to geocentric coordinates, the final step in the calculation of the SSB antenna positions $\mathbf{x}_i(t)$ involves a special relativistic transformation of the positions from geocentric coordinates to SSB coordinates. The earth's orbital motion is obtained from numerically integrated ephemerides developed at the Jet Propulsion Laboratory (Ref. 15).

IV. Propagation Media Calibrations

The total measured delay is corrupted by differential delays due to the propagation media traversed by the radio waves. This section outlines the calibration techniques used to correct for the delays introduced by the troposphere and charged particles.

Due mainly to changes in slant range, the delay produced by the troposphere varies as a function of antenna elevation angle. At a 10-deg elevation angle, the tropospheric delay is approximately 40 nsec (12 m) while, in the local vertical (zenith) direction, it is approximately 7 nsec (2 m). For the present data, a priori corrections for these atmospheric delays were derived from a monthly-mean troposphere model developed at JPL (Chao, Ref. 16) to calibrate radiometric data for spacecraft navigation. For each month of the year, Chao obtained from regional meteorological data mean values for the zenith delays due to the wet and dry components of the troposphere. The tropospheric delay for a given antenna pointing direction can be calculated by mapping the appropriate monthly-mean total zenith delay (dry plus wet) according to an elevation-dependent mapping equation derived from a ray-trace analysis (Ref. 16). Chao estimates the accuracy (1σ) of this calibration technique to be about 3% of the applied correction, or about 40 cm in the worst case. Since the overall uncertainties (1σ) in the delays measured in the later experiments were approximately 15–30 cm, a priori errors in the

tropospheric delays were about the same as the sum of the combined effects of the other observable errors. For this reason, zenith tropospheric delays were also estimated in the multiparameter fit (Sec. VII), but were constrained to the Chao model at the 3% level by means of an additive covariance constraint matrix.

The ionospheric contribution to the observed delays falls approximately in the range 4–50 cm at X-band, depending on antenna elevation and the electron content of the ionosphere. Ionosphere calibrations were handled in two different ways. In later experiments, both S-band and X-band delays were generally measured so that dual-band calibrations were possible. In principle, the ionospheric effect and any other space plasma effect were thereby reduced to a level below other errors in the delay observables in those particular sessions. Ionospheric delays in the single-band sessions were corrected through the use of satellite-Faraday-rotation data. The corrections were obtained by mapping a reference 24-hour zenith delay signature to the time, longitude and, direction of each observation. The reference signature was obtained by averaging the daily Faraday-rotation measurements gathered over a period of 3 months (July–Sept.) in 1971 at Sagamore Hill, Mass. (Ref. 17). It is estimated that this calibration procedure reduces the ionospheric effect by approximately 65%, so that the worst-case observable error due to the ionosphere is roughly 20 cm at X-band. This mean-signature approach to ionosphere calibration was adopted for the single-band experiments due to the unavailability of reliable Faraday-rotation data at the desired times and locations.

V. Observing Strategy

In order to measure both polar motion and Universal Time at a given epoch, concurrent or nearly concurrent measurements on two nonparallel baselines, ideally orthogonal, are required. This requirement is a consequence of the fact that the delay and delay rate observables for a given baseline are insensitive to earth rotations about that baseline vector. In the present experiments, the California/Spain baseline, which is essentially east-west, is sensitive to UT1 and to the X component of polar motion but is relatively insensitive to the Y component of polar motion. In contrast, the California/Australia baseline is sensitive to UT1 and to the Y component of polar motion but is much less sensitive to the X component. Thus, the combination of nearly concurrent observations on both these baselines can yield accurate measurements of all three rotational parameters.

The sources observed on a given day were chosen to cover the full range of directions allowed by the mutual visibility of the two stations. A wide range of source directions is desirable to "separate" sensitivity partial derivatives in the multiparameter fit and thereby to reduce errors in estimated parameters, as well as correlations among the errors. During each observing session, an attempt was made to observe each source at least three times in such a way that the observations completely spanned the interval of mutual visibility between stations for each source. The cycling sequence through the allowed observing directions also was made as random as possible in order to reduce correlations with possible periodic errors in the observables.

To strengthen the overall solution, the source lists for separate sessions were often made identical, particularly when additional measurements could still significantly improve the errors in given source locations. Repetitive observations of the same sources also provide an opportunity for consistency checks, namely whether different subsets of the data yield the same location for a given source or for a given antenna within experimental error. Another requirement affecting source selection was that both intercontinental baselines (California/Spain and California/Australia) observe largely the same sources between −10 and +40 deg declination, the region of the celestial sphere that is mutually visible. The presence of a wide range of sources common to both baselines greatly reduces correlations between parameters, such as the strong correlations found in single-baseline solutions for the California/Australia baseline with its limited mutual visibility. Such a strategy can also remove the near-singularity in declination associated with zero-declination sources in single-baseline solutions with the nearly east-west California/Spain baseline.

VI. Summary of Experiments

Over the last ten years, 48 interferometry sessions have been carried out with eight observing antennas. The diameter and approximate efficiency and zenith system temperature for each of these antennas are presented in Table 1. The first five antennas are DSN facilities; the antenna at Big Pine, Calif., is at the Caltech Owens Valley Radio Observatory, and the Westford, Mass., antenna is at the Haystack Observatory. For each session, Table 2 gives the date, antenna pair, session length, number of sources, number of observations, observable type (delay and/or delay rate), observing frequency, and types of station frequency standard. All together, there were 2382 observations of 117 sources taken during 23 S-band sessions, 5 X-band sessions and 20 dual-band sessions. Depending on the session, the station frequency standards were rubidium, cesium or hydrogen masers.

For each source, Table 3 gives the number of sessions in which that source was observed, the average epoch of observation, and the number of observations of delay and delay rate. Three fourths of the 117 sources are observed at least 20 times, with a high of 161 observations for 4C 39.25. Average observa-

tion epochs (calculated by weighting all observations equally) range from 1976.36 for NRAO 190 to 1980.15 for GC 1128+38.

VII. Fitting Technique

After the observations had been reduced to delay and/or delay rate, the resulting observables were passed to a multiparameter weighted-least-squares program that simultaneously fit all of the delay and delay rate data to obtain estimates of geophysical and astrometric parameters (Ref. 9). Solve-for parameters in the fit could be grouped into two categories. The first category involved parameters specific to a given observing session and included clock rate and epoch, polar motion, Universal Time (UT1) and troposphere parameters for each station. The second category involved "global" parameters common to more than one session and included station locations, two position angles per source, solid-earth-tide parameters, the gamma factor of general relativity and the precession constant. When necessary for any given parameter, the data were organized into segments, over each of which an independent value could be estimated for that parameter in the grand fit to all of the data. For example, this capability was employed for clock parameters, for which it was quite often necessary to divide a session into a number of parts, in each of which a linear or quadratic function of time was applied. To parameterize the troposphere model, the longer sessions were divided into 12-hour portions at each station, each with an independent zenith troposphere delay. Finally, in special tests for repeatability of source-position and station-location parameters, similar divisions were made.

In an ideal measurement of UT/PM with DSN stations, concurrent sessions would be carried out on the California/Spain and California/Australia baselines. Since concurrent sessions are not possible due to constraints of mutual visibility, nearly concurrent sessions on adjacent days were scheduled whenever possible. In this imperfect approach, UT1, which is fairly rapidly varying, was modeled as an independent parameter in each session. On the other hand, more slowly varying polar motion was modeled as a single pole position for the two adjacent sessions. Thus when UT/PM results are presented for adjacent sessions, only a UT1 value will be listed for each session, while PM values will be listed for a fictitious session at the mean time of the two sessions.

The observables in the fit were weighted in inverse proportion to the square of a total error computed as the root-sum-square of the estimated errors from known error sources. The two estimated delay errors were system noise (100 to 545 psec) and ionosphere error for single-band data (35% of ionosphere delay, which equalled 2 to 18 cm for X-band). An explicit troposphere error was not included in the observable weighting since the troposphere was accounted for (to first approximation) through the covariance matrix in the form of a solve-for parameter. In addition to these individually estimated errors, an adjustable session-specific error was root-sum-squared with the other errors in order to account for unestimated or underestimated errors in the data. To estimate the adjustable error, several preliminary fits were made in which this error was adjusted until the normalized chi-square computed separately for the delay and delay rate observables was approximately equal to 1.0 for each session and for the overall fit. Although this treatment of observable noise is far from perfect, it does provide, to a first approximation, an estimate of total observable noise that is based both on a priori information and on the fit residuals.

Two problems encountered in VLBI multiparameter fits are specification of the origin of right ascension and the establishment of an earth-fixed coordinate frame. In this work we have adopted for the right ascension of 3C 273B the value $12^h\ 29^m\ 6\overset{s}{.}6997$ in order to match the definition currently used by many VLBI observing groups. We view this approach as a temporary convenience until it is possible to align our reference frame with the celestial reference frame defined by the JPL planetary ephemeris. To obtain that alignment, the Astronomical Measurements Group at JPL is currently analyzing data from a number of radio interferometry experiments designed to measure the difference in the delay obtained for a spacecraft in orbit about a planet and the delay obtained for a nearby extragalactic source.

For alignment of the axes of the earth-fixed frame with the CIO frame, a two-step procedure based on an a priori covariance matrix is more accurate than the usual approach of relying on BIH values for UT/PM on a single selected reference day. The first step is a fit in which all UT/PM parameters are allowed to vary within the constraints imposed by the BIH values and their errors. This step implicitly minimizes, over the span of the observations, the rms deviation of interferometrically measured UT1 and polar motion from BIH a priori values. In subsequent fits, the UT/PM solve-for values obtained in the first step for two strong adjacent sessions (reference days) are assigned as exactly known quantities. With this approach, subsequent UT/PM values and the axes of the earth-fixed frame are aligned on average with the conventional definitions in the best way provided by the data in hand. For the present data, we estimate that the resulting alignment with BIH conventions is accurate to 2–5 mas for the polar axis and to 0.4–1.0 msec for longitude.

Since interferometry observables for extragalactic sources are insensitive to the origin of the earth-fixed frame, data content by itself does not require the consideration of that

origin. However, for computational reasons, it is desirable to use station locations as the solve-for parameters rather than baseline vectors. With this approach, at least one station (the reference station) must be assigned an a priori position to establish the origin of the earth-fixed frame and thereby prevent singularity in the least-squares solution. Thus, in addition to the axis orientation error mentioned above, the adjusted locations of the other stations are subject to errors that are a consequence of errors in the a priori location of the reference station.

For the present data, the choice of a reference station is relatively simple. Of the eight stations at which the observations were made (see Table 1), all but two are linked to the deep space station (DSS 14) at Goldstone, Calif., either by direct two-station observations or through a common third station. The two exceptions are the Owens Valley (OVRO) and Westford, Mass. (HAYST) antennas, for which a disjoint 48-hour session comprises the only data. To avoid defining a second reference station, the location of OVRO was linked to DSS 13 at Goldstone by employing DSS 13–OVRO baseline measurements made by the ARIES project (Ref. 18). In our fits, each component of this baseline was constrained to equal the ARIES result on the basis of the ARIES formal uncertainties (~9 cm). The only other sessions not involving station DSS 14 (DSS 11–DSS 43 on 77/2/1 and DSS 62–DSS 63 on 79/11/15) have both stations participating in direct observations with DSS 14 on other dates. Thus, with the addition of this constraint on the DSS 13–OVRO baseline, all seven station locations could be referenced to an assigned value for the DSS 14 location. For this assigned value, we chose the coordinates from station location set LS111A derived from spacecraft tracking (Ref. 19). Moyer estimates the DSS 14 location error to be ~1 m in spin radius, ~2–5 m in longitude, and ~10 m in Z-height. As mentioned above, these reference station errors will lead to a systematic bias in all station positions. On the other hand, the relative locations of the stations have uncertainties that are determined by the quality of the VLBI observations, the accuracy of the delay model, and the size of the alignment errors for the earth-fixed axes.

VIII. Results

The majority of the results presented in this section were derived from a single "standard" fit to the 1971–80 delay and delay-rate observables in which parameters were adjusted for source positions, station locations and clocks, UT/PM, tropospheric delays, gravitational bending and solid earth tides. In addition to the features discussed in Sec. VII, astronomical constants, time scales and the fundamental reference frame were in accord with the IAU resolutions to be implemented to 1984 (as summarized by Kaplan (Ref. 10)).

For an overall view of the goodness of fit, Table 4 shows for each session the rms residual delay and delay rate obtained with the standard fit. While the overall averages are 0.52 nsec and 0.30 psec/sec, variation among sessions is considerable, with the best sessions giving substantially smaller residuals.

To assess the contribution of the delay rate observables to the adjusted parameters and their estimated uncertainties, another fit was performed. This was identical to the standard fit, except that only the delay observables were used, and thus only sessions after 77/1/12 were included (see Table 2). The resulting parameter values were consistent with those from the standard fit. Typically, formal uncertainties from the standard fit were smaller than those from the delay-only fit by 30% for station coordinates, 20% for source positions, and 15% for Universal Time and polar motion. Numerous fits were necessary to determine the magnitudes of the adjustable observable errors (see Sec. VII) that corresponded to chi-square values of 1.0. These fits indicated that all solve-for parameters were fairly insensitive to the values of the adjustable errors. For example, any given solve-for parameter varied by a small fraction of its formal uncertainty when the adjustable errors were varied over a range corresponding to chi-square between 0.9 and 1.1. Additional special fits to the data have been performed for specific purposes such as source-position and baseline repeatability tests and adjustment of nutation and precession, as discussed below.

A. Source Positions

The last columns of Table 3 give, in terms of right ascension and declination, the J2000.0 positions of 117 sources resulting from the standard fit to the 1971–80 VLBI data. Of these, seven (3C 48, 3C 119, 3C 138, 3C 309.1, 3C 395, 3C 418 and DA 611) were observed only on short baselines, and thus have formal position errors exceeding 0.″1. One source (4C 55.16) was observed only once on an intercontinental baseline and its position must therefore be used with caution. Another, low-declination source (P 1130 + 009) was not observed on the California/Australia baseline and therefore has a large declination error. Other sources for which there are fewer than 5 delay observations may also lack adequate redundancy. All listed errors are 1σ formal uncertainties obtained from the covariance matrix of the standard fit and thus do not properly account for errors due to mismodeling of precession and nutation.

The statistics of our source catalog are more clearly demonstrated in Figs. 1–3. Figure 1 presents the source distribution in terms of right ascension and declination, with error bars indicating formal uncertainties. (Note the different scales for position and uncertainty.) In Figs. 2 and 3, histograms of the arc length errors for RA and declination indicate that the

mean position error is approximately 5 mas for both coordinates.

In order to determine the repeatability of source positions over a long period of time, and to test the validity of the formal error estimates, the data were divided into two segments: observations made before January 1979 and observations made after May 1979. The average measurement epochs in these two segments were approximately 1977.5 and 1979.9. All sources with more than 10 (delay + delay rate) observations in each segment were assigned independent position parameters in the two segments. A fit was performed in which all other parameters were solved for as in the standard fit. In Table 5, columns 4 and 5 show the RA and declination differences for the two position estimates, columns 6 and 7 give the root-sum-squared errors, and the last two columns give the differences normalized by the rss errors. The rms differences for right ascension and declination (2.9 and 4.0 mas) are somewhat smaller than the average formal uncertainties for the catalog of Table 3. Figure 4 shows a plot of the position differences in columns 4 and 5 of Table 5. Since Table 5 shows that no normalized difference exceeds 1.6σ, the positions produced by the two segments of data are in very good agreement at the level of the formal uncertainties. Overall, the results indicate that our formal source position errors are close to or perhaps slightly larger than the true random errors. Preliminary comparison with the source catalog of another VLBI group (Refs. 20, 21) resulted in fair agreement but indicated that about a 50% increase in formal uncertainties might be needed.

B. Station Locations and Baselines

Table 6 presents the station locations obtained from the standard fit to the entire data span. As discussed in Sec. VII, station DSS 14 at Goldstone is the reference station, and the other seven locations are adjusted. The value 299792.458 km/sec was used for the velocity of light. Formal error estimates of the station coordinates range from ~2 cm for DSS 13 to ~56 cm for the poorly determined equatorial component of DSS 11. The formal uncertainties for the station locations are relative errors and do not account for the uncertainties from orienting the axes of the earth-fixed frame or from the assignment of the DSS 14 location.

Rather than present all 28 baselines between the eight stations, we focus attention on the six typical station pairs given in Table 7, which include two intercontinental (DSS 14 to DSS 43 and DSS 63), one transcontinental (OVRO to HAYST), and three local (DSS 14 to DSS 11 and DSS 13, DSS 62 to DSS 63) baselines. Anomalously high formal errors for the DSS 14–DSS 11 baseline are due to the fact that no direct DSS 14–DSS 11 observations were included in the fit. Since baseline errors, unlike baseline vectors, cannot be computed accurately from Table 7, formal uncertainties of components and lengths of all 28 baselines are listed in Table 8. These formal uncertainties do not account for the errors in orienting the axes of the earth-fixed frame.

Two comparisons of baseline lengths with results of independent investigations are possible. For the OVRO-HAYST baseline, the east coast VLBI group (Ref. 22) obtained a length of 3928881.59 ± 0.02 m based on data extending from July to October 1980. This result is in excellent agreement with our length measurement. A 1975 ground survey of the DSS 62–DSS 63 baseline (Ref. 23) gives a length of 10452.61 m, also in excellent agreement with our result.

To test for baseline length repeatability, another special fit to all the data was performed in which independent station location parameters were assigned to DSS 43 and DSS 63 for each relevant post–1976 intercontinental session of Table 2. This led to 20 estimated values for the Australian station (DSS 43) and 12 for the Spanish station (DSS 63), with both spanning a 3-year period. The Goldstone station (DSS 14) was assigned the same common reference location (see Table 6). To avoid a singular solution, the UT/PM parameters were held fixed at the results of the standard fit. All other parameters were treated as in the standard fit. The baseline length results from this fit, which are shown in Figs. 5 and 6 along with 1σ formal uncertainties, exhibit several interesting features. First, we note the general decrease of the error estimates with time, which indicates continuing improvement in system performance. Second, comparison of X-band and S/X results with S-band results indicates that the S-band baselines were corrupted by as much as 100 cm by ionospheric effects for both baselines. Third, even though there may be a suggestion of change, there is no convincing evidence of relative motion between stations. To quantitatively assess possible changes in baseline lengths over the years, however, a linear function of time has been separately fit to the length results for each of these two intercontinental baselines. Due to ionospheric corruption, only the X-band and dual-band results were included in these fits, which led to slopes of -5 ± 5 and $+7 \pm 10$ cm/yr for the baselines to Australia and Spain, respectively. It is of interest to note that, even though these results are consistent with no motion over the 3-year period, both slopes are closer to the length changes inferred (Ref. 24) from a global model (Ref. 25) of plate motion over geological time scales than they are to the assumption of no motion. Morabito's application of the Minster-Jordan model indicates length changes of approximately –4 and +2.5 cm/yr for the baselines to Australia and Spain, respectively. Doppler satellite tracking (Ref. 26) has recently produced measurements of plate motion which appear to be consistent with our baseline results.

C. Polar Motion and UT1

Table 9 presents all the UT/PM values obtained from the standard fit, while Figs. 7 and 8 compare our X and Y polar motion results for 1977–80 with BIH Circular D (Ref. 27). Our UT/PM uncertainties are relative errors and do not include errors resulting from orienting the axes of the earth-fixed frame. If the BIH PM values are assigned an uncertainty of approximately 0.″01, there are no outstanding discrepancies between the two techniques. For the UT1 results similarly plotted in Fig. 9, short-period tidal fluctuations have been removed from the VLBI data in order to permit comparison with the heavily smoothed BIH values. The solid curve represents lunar laser ranging (LLR) data as smoothed over a 10-day interval by Fliegel et al. (Ref. 28) with the same tidal dependence removed. These lunar data originally consisted of several hundred points in the range of the plot. The LLR curve has been displaced vertically in order to remove a 1.0-msec bias between VLBI and LLR values. The figure shows that the UT1 values measured by VLBI, LLR, and BIH generally agree with one another if the BIH values are assigned errors of approximately 2 msec, and the LLR values errors of 1 msec or less. Both the VLBI and LLR results suggest the same oscillation of ~2 msec amplitude about the BIH values. Recent work by Capitaine and Feissel (Ref. 29) indicates that introduction of the 1980 IAU nutation model into the BIH solution results in corrections to UT1-UTC as large as 2 msec. We plan to investigate the impact of such a correction on the above comparison of UT1 results. Two of the three points of large discrepancy (~2 to 3σ) between VLBI and LLR (in February 1980) occurred at the center of a 20-day gap in the LLR data. Thus these differences may represent a real short-period excursion that was not sampled in the LLR measurements. The third point of large discrepancy in February 1977 remains unexplained.

D. Global Parameters

In addition to the source positions and station locations discussed above, global parameters include the precession constant, earth-tide parameters, and the gamma factor of general relativity. The long time span of the data provided an opportunity to solve for the precession constant. Specifically, if we solve for a residual precession rate of the earth, we find that the best fit to our data does not occur for the 1976 IAU precession constant, but for a value smaller than that value by 3.8 ± 0.9 mas/yr (for luni-solar precession). However, this result must be qualified by the observation that the first 6.5 years of our 8.5-yr data span consisted of less reliable data (S-band delay rate only and S-band delay and delay rate data) which may have biased our solution. Since our data cover a time interval small compared to the 18.6-year nutation period, there is a high degree of correlation between the precession constant and the nutation amplitude with that period. Thus our present data cannot accurately separate these two effects. Such a separation will become possible, however, as the span of data approaches a significant fraction of the 18.6-year period. Similar explorations of modified nutation and precession by the JPL group analyzing lunar laser ranging (Ref. 30) have yielded estimates of 8 ± 8 mas/yr for the 18.6-yr nutation term correction, and −2 ± 4 mas/yr for the precession correction.

Since the data were inadequate for determination of independent earth-tide parameters at each station, a universal set of parameters was specified for all stations. This model allowed adjustment of three parameters: the vertical and horizontal Love numbers and the tide phase lag. The standard fit yielded values of 0.63 ± 0.03 and 0.058 ± 0.016 for the vertical and horizontal Love numbers, and 0.0 deg ± 1.5 deg for the tide phase lag. These results are in good agreement with the commonly accepted values of 0.603–0.611, 0.0832–0.0842 and 0 deg (Refs. 31, 32).

Since gravitational deflection of the incoming signal by the sun is a relatively large effect for these long baselines even at large sun-earth-source angles, it is possible to solve for the gamma factor of the parameterized post-Newtonian formalism. The result of 0.997 ± 0.041 is in good agreement with general relativity (Ref. 33).

E. Troposphere Parameters

The dry troposphere parameters obtained from the standard fit are presented in Table 10 along with 1σ formal error estimates. As discussed in Sec. IV, the tropospheric delays for the DSN stations were constrained to the Chao model on the basis of the estimated error (3%) in that model. For sessions involving delay observations (1977–80), the formal uncertainties from our fit are considerably better than the Chao a priori errors. In addition to the standard fit, which was based on 12-hour subdivisions of tropospheric delays, another fit was made in which new troposphere parameters were introduced at each station at 6 a.m. and 6 p.m. local time, in order to investigate the effect of dividing tropospheric delays into day and night portions. No significant effects on any of the adjusted parameters were observed in this fit. The possibility of mismodeling the tropospheric delays at low elevation angles was also considered. A fit that omitted any observation with an elevation angle less than 10 deg at either station produced parameters that differed from those given by the standard fit by no more than the formal uncertainties. Attempts to find correlations between the delays in Table 10 and local atmospheric measurements (when available) were unsuccessful.

F. Clock Parameters

As shown in Table 11, 170 station clock parameters were adjusted in the standard fit to describe clock epoch and rate

offsets and nonlinearities in the delay data. The reference clock is that at DSS 14, with three exceptions. In the two three-station experiments in February 1977, it was necessary to use one of the other station clocks (DSS 11 or DSS 43) as a reference, because of missing data at DSS 14. For the OVRO-HAYST observations in February 1978, the Haystack clock served as a reference. A quadratic clock model was used in only three sessions, while piecewise linear fits were required in 23 of the 48 sessions. For the best sessions, the offsets between station clocks were measured with formal uncertainties of the order of 0.01 psec/sec for clock rate and 0.5 nsec for clock epoch. Since instrumental phase calibration was not employed, the "solve-for" clock parameters are not purely clock offsets, but also include instrumental terms. Thus these accuracies are merely indicative of the capability of this technique once proper instrument calibration is performed. The 73 clock parameters adjusted for the delay rate observables are not reported here.

IX. Discussion and Conclusions

The development of two radio interferometry systems at JPL during the past decade has led to a number of significant astrometric and geophysical results. With the present system, we have measured the lengths of two DSN baselines (California/Australia and California/Spain) with formal uncertainties of 10 cm and 4 cm, respectively. When the baselines were adjusted independently for each observing session, no convincing evidence of change in the length of either baseline was detected. We are continuing to improve the system and accumulate data. If actual rates of tectonic displacement are of the order of 5 cm/yr, we expect to detect them within a few years.

A radio source reference frame has been established which contains 104 sources with positional uncertainties of approximately 5 mas. Our estimates for parameters describing earth tides and gravitational bending agree with commonly accepted values. Further, formal uncertainties in adjusted "clock" parameters indicate a potential for synchronizing station clocks at the nanosecond level.

Our measurements of polar motion and UT1-UTC with formal uncertainties of 5 to 20 cm agreed fairly well with results obtained by the BIH and lunar laser ranging. A fit to our current data produced an estimate for the luni-solar precession constant which is smaller than the IAU value by 3.8 ± 0.9 mas/yr. Because our earlier data are less reliable, more observations are required to verify this result. Since the discrepancy of 3.8 mas/yr is substantially larger than the estimated uncertainty in the IAU value (1 or 2 mas/yr), there is strong motivation to improve the reliability of this result through more measurements.

The results of these observations underscore the importance of improving the scope and accuracy of radio interferometry measurements in a number of areas.

Acknowledgments

J. G. Williams and P. F. MacDoran made pioneering contributions to the radio interferometry program at JPL. We thank the many staff members at the DSS 11, 13, 14, 43, 62, 63, OVRO, and Haystack antennas for their expert attention to instrumentation at the observing stations. A. Rius planned and executed the DSS 62–DSS 63 experiment in November 1979. Finally, we thank J. D. Gunckel for her many hours of dedication in seeing the data through the initial stages of processing.

References

1. Thomas, J. B., et al., "A Demonstration of an Independent-Station Radio Interferometry System with 4-cm Precision on a 16-km Baseline," *J. Geophys. Res.*, 81, pp. 995–1005, 1976.

2. Rogers, A. E. E., et al., "Geodesy by Radio Interferometry: Determination of a 1.24-km Base Line Vector With ~5-mm Repeatability," *J. Geophys. Res.*, 83, pp. 325–334, 1978.

3. Ryan, J. W., et al., "Precision Surveying Using Radio Interferometry," *J. Surveying and Mapping Div., Proc. ASCE*, Vol. 104, No. SU1, pp. 25-34, 1978.

4. Niell, A. E., et al., "Comparison of a Radio Interferometric Differential Baseline Measurement With Conventional Geodesy," *Tectonophysics*, 52, pp. 49-58, 1979.

5. Herring, T. A., et al., "Geodesy by Radio Interferometry: Intercontinental Distance Determinations with Subdecimeter Precision," *J. Geophys. Res.*, 86, pp. 1647-1651, 1981.

6. Thomas, J. B., *An Analysis of Long Baseline Radio Interferometry*, Technical Report 32-1526; Vol. VII, 37-50; Vol. VIII, 29-38; Vol. XVI, 47-64, Jet Propulsion Laboratory, Pasadena, Calif., 1972.

7. Thomas, J. B., *An Analysis of Radio Interferometry With the Block 0 System*, Publication 81-49, Jet Propulsion Laboratory, Pasadena, Calif., 1981.

8. Rogers, A. E. E., "Very-Long-Baseline Interferometry with Large Effective Bandwidth for Phase-Delay Measurements," *Radio Science*, 5, 1239-1247, 1970.

9. Fanselow, J. L., "Observation Model and Parameter Partials for the JPL VLBI Parameter Estimation Software MASTERFIT-V1.0," Jet Propulsion Laboratory, Pasadena, Calif., 1983 (to be published).

10. Kaplan, G. H., "The IAU Resolutions on Astronomical Constants, Time Scales, and the Fundamental Reference Frame," USNO Circular No. 163, United States Naval Observatory, Washington, D.C., 1981.

11. Lieske, J. H., et al., "Expressions for the Precession Quantities Based Upon the IAU (1976) System of Astronomical Constants," *Astron. Astrophys.* 58, pp. 1-16, 1977.

12. Wahr, J. M., "The Forced Nutations of an Elliptical, Rotating, Elastic and Oceanless Earth," *Geophys. J. Roy. Astron. Soc.*, 64, 705, 1981.

13. Seidelmann, P. K., "1980 IAU Theory of Nutation: The Final Report of the IAU Working Group on Nutation," *Celestial Mechanics*, 27, pp. 79-106, 1982.

14. Aoki, S., et al., "The New Definition of Universal Time," *Astron. Astrophys.* 105, pp. 359-361, 1982.

15. Standish, E. M., "Orientation of the JPL Ephemerides, DE200/LE200, to the Dynamical Equinox of J2000," *Astron. Astrophys.*, 114, pp. 297-302, 1982.

16. Chao, C. C., *The Tropospheric Calibration Model for Mariner Mars 1971*, Technical Report 32-1587, pp. 61-76, Jet Propulsion Laboratory, Pasadena, Calif., 1974.

17. Klobuchar, J. A., and Malik, C. A., *Geophysics and Space Data Bulletin*, Vol. 9, No. 2, pp. 311-318, A. L. Carrigan, ed.; AFCRL-72-0502, Special Report No. 145, Air Force Cambridge Research Laboratories, L. G. Hanscom Field, Mass., 1972.

18. Wallace, K. S., private communication, 1982.

19. Moyer, T. D., private communication, 1981.

20. Ma, C., Clark, T. A., and Shaffer, D. B., Mark III VLBI: Astrometry and Epoch J2000.0, Bull. AAS 13, 899, 1982.

21. Ma, C., private communication, 1982.

22. Rogers, A. E. E., et al., "Very-Long-Baseline Interferometry: The Mark III System for Geodesy, Astrometry and Aperture Synthesis," *Science*, 219, pp. 51-54, 1983.

23. Rius, A. and Calera, E., private communication, 1982.

24. Morabito, D. D., Claflin, E. S., and Steinberg, C. J., "VLBI Detection of Crustal Plate Motion Using DSN Antennas as Base Stations," *DSN Progress Report 42-56*, pp. 59-75 Jet Propulsion Laboratory, Pasadena, Calif., 1980.

25. Minster, J. B., and Jordan, T. H., "Present-Day Plate Motions," *J. Geophys. Res.*, 83, pp. 5331-5354, 1978.

26. Anderle, R. J., and Malyevac, C. A., "Current Plate Motions Based on Doppler Satellite Observations," *Geophys. Res. Lett.*, 10, p. 67, 1983.

27. Bureau International de l'Heure, *Annual Report for 1981*, Paris, France, 1981.

28. Fliegel, H. F., Dickey, J. O., and Williams, J. G., "Intercomparison of Lunar Laser and Traditional Determinations of Earth Rotation," pp. 53-88 *IAU Colloquium Proc.* No. 63, O. Calame, ed., Grasse, France, 1981.

29. Capitaine, N., and Feissel, M., "The Introduction of the IAU 1980 Nutation Theory in the Computation of the Earth Rotation Parameters by the Bureau International de l'Heure, a Proposal," *Bull. Geodesique*, 1982 (to be published).

30. Dickey, J. O., and Williams, J. G., "Geophysical Applications of Lunar Laser Ranging," *Eos*, 63, 301, 1982.

31. Wahr, J. M., "The Tidal Motions of a Rotating, Elastic and Oceanless Earth," pp. 162-171, Thesis, University of Colorado, 1977.

32. Lambeck, K., *The Earth's Variable Rotation: Geophysical Causes and Consequences*, pp. 13, 111, Cambridge, Cambridge U. Press, 1980.

33. Will, C. M., "The Confrontation Between General Relativity and Experiment," *Ann. N.Y. Acad. Sci.*, 336, pp. 307-321, 1980.

Table 1. Characteristics of antennas used in 1971–1980 VLBI observations

Antenna	Location	Diameter (m)	Efficiency		Zenith system temperature (K)	
			S	X	S	X
DSS11	Goldstone, Calif.	26	0.55	-	35	-
DSS13	Goldstone, Calif.	26	0.58	0.45	35	35
DSS14	Goldstone, Calif.	64	0.56	0.43	25	25
DSS43	Tidbinbilla, Australia	64	0.56	0.43	25	25
DSS62	Madrid, Spain	26	0.58	-	35	-
DSS63	Madrid, Spain	64	0.56	0.43	25	25
OVRO	Big Pine, Calif.	40	-	0.45	-	200
HAYST	Westford, Mass.	37	-	0.40	-	75

Table 2. Summary of 1971–1980 VLBI observing sessions

Date	Antennas	Session length (hr)	No. of distinct sources	No. of obs.	Observable type *	Freq. bands	Freq. standard **
71/ 8/28	14-62	15	17	43	DR	S	H,H
71/ 9/ 1	14-62	15	13	24	DR	S	H,H
71/ 9/ 6	14-62	17.5	15	45	DR	S	H,H
71/ 9/10	14-62	15	15	45	DR	S	H,H
73/ 4/30	14-62	8	12	21	DR	S	H,Rb
73/ 9/ 8	14-62	8	7	17	DR	S	H,Rb
74/ 2/15	14-62	6.5	8	20	DR	S	H,Rb
74/ 4/21	14-62	8	10	20	DR	S	H,Rb
74/ 6/21	14-62	6	10	17	DR	S	H,Rb
74/ 8/ 6	14-62	5	3	7	DR	S	H,Rb
77/ 1/12	11-43	2	9	12	D+DR	S	Rb,H
77/ 1/21	11-14-43	4.5	20	43	D+DR	S	Rb,H,H
77/ 1/31	14-63	6	15	26	D+DR	S	H,Rb
77/ 2/ 1	11-43	5.5	16	23	D+DR	S	Rb,H
77/ 2/13	11-14-43	10	24	68	D+DR	S	Rb,H,H
77/ 2/28	11-14-43	10	25	64	D+DR	S & X	Rb,H,H
77/ 4/13	14-63	7.5	19	44	D+DR	S	H,Rb
78/ 1/14	14-43	14	24	50	D+DR	X	H,H
78/ 1/24	14-43	6	19	36	D+DR	X	H,Rb
78/ 2/12	14-43	8	18	44	D+DR	X	H,H
78/ 2/24	OV-HY	48	8	123	D+DR	X	H,H
78/ 5/15	14-43	10	23	58	D+DR	X	Cs,H
78/ 7/30	14-63	9	25	47	D+DR	S	H,H
78/ 9/ 3	14-43	5	19	29	D+DR	S/X	H,Rb
78/ 9/ 4	14-63	9	13	25	D+DR	S	H,H
78/10/27	14-43	18.5	45	97	D+DR	S/X	H,H
78/10/30	14-63	8	22	32	D+DR	S/X	H,Cs
78/11/ 4	14-43	3.5	15	23	D+DR	S/X	H,H
78/11/ 5	14-63	23	30	60	D+DR	S/X	H,Cs
78/12/31	14-43	22.5	48	117	D+DR	S/X	H,H
79/ 6/ 6	13-14	3.5	13	15	D+DR	S	H,H
79/ 7/21	13-14	4.5	23	30	D+DR	S	H,H
79/ 8/26	13-14	6	24	31	D+DR	S	H,H
79/ 9/18	13-14	5	22	35	D+DR	S	H,H
79/11/15	62-63	4.5	16	25	D+DR	S	Rb,H
79/11/23	14-43	9	21	39	D+DR	S/X	H,H
79/11/25	14-63	20.5	46	113	D+DR	S/X	H,H
79/12/20	14-43	23	60	137	D+DR	S/X	H,H
79/12/21	14-63	10	16	33	D+DR	S/X	H,H
79/12/27	14-63	23	30	66	D+DR	S/X	H,H
79/12/29	14-43	8.5	21	45	D+DR	S/X	H,H
80/ 1/12	14-43	20	52	110	D+DR	S/X	H,H
80/ 1/25	14-63	20.5	26	47	D+DR	S/X	H,H
80/ 1/27	14-43	8.5	22	42	D+DR	S/X	H,H
80/ 2/13	14-63	23	52	137	D+DR	S/X	H,H
80/ 2/14	14-43	8.5	25	50	D+DR	S/X	H,H
80/ 2/23	14-43	20.5	54	99	D+DR	S/X	H,H
80/ 2/24	14-63	8.5	24	48	D+DR	S/X	H,H
Total			117	2382			

* D = delay, DR = delay rate.
** Rb = rubidium, Cs = cesium, H = hydrogen maser.

Table 3. JPL 1983–2 radio source catalog (J2000.0)

	IAU designation	Common name	Average obs. epoch	No. of sessions	Observations Delay	Observations Delay rate	Right ascension h	m	s	Error(σ)	Declination d	m	s	Error(σ)
	0008-264	P 0008-264	1980.05	3	8	8	0	11	1.24747	0.00039	-26	12	33.3891	0.0053
	0104-408	P 0104-408	1979.43	7	21	21	1	6	45.10824	0.00023	-40	34	19.9640	0.0035
	0106+013	P 0106+01	1978.39	23	38	46	1	8	38.77105	0.00011	1	35	0.3174	0.0026
	0111+021	P 0111+021	1980.04	3	6	6	1	13	43.14505	0.00168	2	22	17.3155	0.0254
	0113-118	P 0113-118	1978.85	5	12	12	1	16	12.52214	0.00067	-11	36	15.4375	0.0103
	0133+476	DA 55	1979.14	20	41	43	1	36	58.59495	0.00018	47	51	29.1053	0.0013
S	0134+329	3C 48	1979.64	3	3	3	1	37	41.34210	0.05828	33	9	35.4968	0.3769
	0202+149	P 0202+14	1980.02	6	13	13	2	4	50.41405	0.00012	15	14	11.0461	0.0024
	0224+671	DW 0224+67	1977.48	24	39	62	2	28	50.05199	0.00032	67	21	3.0339	0.0019
	0234+285	CTD 20	1980.02	6	13	13	2	37	52.40585	0.00014	28	48	8.9942	0.0021
	0235+164	GC 0235+16	1980.04	6	11	11	2	38	38.93026	0.00012	16	36	59.2767	0.0023
R	0237-233	P 0237-23	1980.02	2	3	3	2	40	8.17472	0.00091	-23	9	15.7416	0.0126
	0300+470	OE 400	1978.39	10	19	19	3	3	35.24246	0.00024	47	16	16.2831	0.0019
	0316+413	3C 84	1978.72	10	11	16	3	19	48.16043	0.00019	41	30	42.1074	0.0019
	0332-403	P 0332-403	1977.84	5	17	17	3	34	13.65384	0.00037	-40	8	25.3969	0.0044
	0333+321	NRAO 140	1977.61	25	34	50	3	36	30.10794	0.00016	32	18	29.3448	0.0017
	0336-019	CTA 26	1979.52	9	13	15	3	39	30.93773	0.00013	-1	46	35.8004	0.0029
	0355+508	NRAO 150	1978.60	6	26	26	3	59	29.74801	0.00026	50	57	50.1688	0.0022
	0402-362	P 0402-362	1978.07	4	11	11	4	3	53.74941	0.00044	-36	5	1.9132	0.0054
	0406+121	GC 0406+12	1980.03	6	12	12	4	9	22.00860	0.00014	12	17	39.8426	0.0055
	0420-014	P 0420-01	1977.97	11	19	20	4	23	15.80052	0.00020	-1	20	33.0642	0.0037
	0420+417	VRO 41.04.01	1979.72	7	9	10	4	23	56.01004	0.00026	41	50	2.7185	0.0033
S	0429+415	3C 119	1979.66	3	3	2	4	32	36.48682	0.28126	41	38	28.7574	1.6483
	0430+052	3C 120	1977.28	15	16	24	4	33	11.09592	0.00025	5	21	15.6142	0.0040
	0434-188	P 0434-188	1980.02	3	5	5	4	37	1.48289	0.00057	-18	44	48.6198	0.0086
	0438-436	P 0438-43	1977.87	5	20	20	4	40	17.17941	0.00033	-43	33	8.6054	0.0040
	0440-003	NRAO 190	1976.36	19	16	34	4	42	38.66089	0.00051	-0	17	43.4243	0.0080
	0451-282	P 0451-28	1980.02	3	5	5	4	53	14.64585	0.00048	-28	7	37.3195	0.0062
S	0518+165	3C 138	1979.55	1	2	2	5	21	10.04094	0.11462	16	38	21.9327	0.5140
	0528+134	P 0528+134	1980.03	7	14	14	5	30	56.41692	0.00012	13	31	55.1501	0.0022
	0537-441	P 0537-441	1978.46	9	20	20	5	38	50.36073	0.00028	-44	5	8.9386	0.0037
	0552+398	DA 193	1977.55	17	27	33	5	55	30.80616	0.00022	39	48	49.1667	0.0019
	0605-085	P 0605-08	1978.96	4	6	6	6	7	59.69940	0.00070	-8	34	49.9881	0.0100
	0607-157	P 0607-15	1978.52	12	18	18	6	9	40.94952	0.00025	-15	42	40.6778	0.0044
	0723-008	DW 0723-00	1979.99	5	8	8	7	25	50.63977	0.00020	-0	54	56.5434	0.0038
	0727-115	P 0727-11	1977.93	12	20	21	7	30	19.11287	0.00024	-11	41	12.6140	0.0043
	0735+178	P 0735+17	1977.85	16	20	24	7	38	7.39401	0.00022	17	42	18.9934	0.0036
	0738+313	OI 363	1977.01	14	14	23	7	41	10.70377	0.00043	31	12	0.2264	0.0054
	0742+103	DW 0742+10	1978.21	22	45	49	7	45	33.05953	0.00013	10	11	12.6897	0.0023
	0748+126	P 0748+126	1980.00	5	10	10	7	50	52.04562	0.00017	12	31	4.8263	0.0031
	0814+425	OJ 425	1977.83	19	23	27	8	18	16.00000	0.00023	42	22	45.4144	0.0018
	0823+033	P 0823+033	1980.01	8	13	13	8	25	50.33852	0.00018	3	9	24.5130	0.0033
	0827+243	DZ 0827+24	1979.76	8	10	11	8	30	52.08663	0.00034	24	10	59.8105	0.0044
R	0831+557	4C 55.16	1977.58	6	4	7	8	34	54.90177	0.00182	55	34	21.1328	0.0228
	0836+710	4C 71.07	1979.64	8	13	15	8	41	24.36807	0.00043	70	53	42.1772	0.0020
	0851+202	OJ 287	1978.04	15	37	41	8	54	48.87506	0.00013	20	6	30.6363	0.0018
	0859-140	P 0859-14	1977.80	11	16	16	9	2	16.83076	0.00081	-14	15	30.8847	0.0123
	0859+470	OJ 499	1977.92	5	6	6	9	3	3.99136	0.00076	46	51	4.1266	0.0051
	0923+392	4C 39.25	1978.52	29	78	83	9	27	3.01394	0.00013	39	2	20.8497	0.0012
	0952+179	AO 0952+17	1979.98	5	5	5	9	54	56.82357	0.00078	17	43	31.2228	0.0126
	1004+141	GC 1004+14	1980.01	6	16	15	10	7	41.49848	0.00047	13	56	29.5911	0.0063
	1034-293	P 1034-293	1980.00	6	10	10	10	37	16.07921	0.00030	-29	34	2.8183	0.0042
	1038+064	OL 064.5	1980.01	2	6	6	10	41	17.16237	0.00020	6	10	16.9218	0.0382
	1040+123	3C 245	1980.03	6	7	6	10	42	44.60596	0.00036	12	3	31.2536	0.0049
	1055+018	P 1055+01	1979.73	10	23	25	10	58	29.60516	0.00011	1	33	58.8174	0.0025
	1104-445	P 1104-445	1977.74	10	15	15	11	7	8.69332	0.00047	-44	49	7.6227	0.0047
	1111+149	GC 1111+14	1980.03	6	11	11	11	13	58.69532	0.00048	14	42	26.9449	0.0061
R	1116+128	P 1116+12	1980.11	3	4	4	11	18	57.30154	0.00031	12	34	41.7093	0.0064

Table 3 (contd)

	IAU designation	Common name	Average obs. epoch	No. of sessions	Observations Delay	Observations Delay rate	Right ascension h	m	s	Error(σ)	Declination d	m	s	Error(σ)
	1123+264	P 1123+26	1980.02	8	21	21	11	25	53.71196	0.00019	26	10	19.9738	0.0023
	1127-145	P 1127-14	1978.70	22	51	51	11	30	7.05233	0.00016	-14	49	27.3941	0.0033
R	1128+385	GC 1128+38	1980.15	1	3	3	11	30	53.28215	0.00047	38	15	18.5471	0.0037
R	1130+009	P 1130+009	1980.05	3	3	3	11	33	20.05385	0.00741	0	40	53.4607	2.2108
	1144-379	P 1144-379	1978.73	15	40	40	11	47	1.37021	0.00022	-38	12	11.0305	0.0034
	1148-001	P 1148-00	1978.09	10	15	15	11	50	43.87056	0.00040	-0	23	54.2102	0.0064
	1222+037	P 1222+037	1980.03	7	13	13	12	24	52.42221	0.00021	3	30	50.2008	0.0038
	1226+023	3C 273	1978.62	20	54	57	12	29	6.69970	0.00000	2	3	8.5914	0.0025
	1228+126	3C 274	1980.04	7	12	12	12	30	49.42342	0.00031	12	23	28.0376	0.0043
	1244-255	P 1244-255	1980.02	6	16	16	12	46	46.80184	0.00025	-25	47	49.2957	0.0037
	1253-055	3C 279	1978.24	6	14	14	12	56	11.16647	0.00038	-5	47	21.5321	0.0063
	1308+326	B2 1308+32	1980.02	10	24	24	13	10	28.66381	0.00013	32	20	43.7788	0.0015
	1313-333	OP-322	1978.15	9	16	16	13	16	7.98528	0.00031	-33	38	59.1782	0.0041
	1334-127	DW 1335-12	1977.96	10	20	21	13	37	39.78288	0.00028	-12	57	24.7031	0.0043
	1342+663	GC 1342+663	1980.12	3	5	5	13	44	8.67956	0.00053	66	6	11.6365	0.0025
R	1349-439	P 1349-439	1980.11	2	4	4	13	52	56.53523	0.00061	-44	12	40.4008	0.0057
	1354+195	P 1354+19	1980.02	11	22	22	13	57	4.43660	0.00017	19	19	7.3665	0.0027
	1418+546	GC 1418+54	1980.13	2	6	6	14	19	46.59754	0.00019	54	23	14.7823	0.0015
	1430-178	OQ-151	1980.05	4	9	9	14	32	57.68946	0.00054	-18	1	35.2438	0.0080
S	1458+718	3C 309 1	1979.61	4	7	7	14	59	7.33716	0.12838	71	40	20.4458	0.6309
	1502+106	OR 103	1979.35	18	39	40	15	4	24.97966	0.00011	10	29	39.1945	0.0023
	1510-089	P 1510-08	1978.18	6	15	15	15	12	50.53332	0.00051	-9	5	59.8409	0.0079
	1519-273	P 1519-273	1980.03	6	15	15	15	22	37.67552	0.00022	-27	30	10.7889	0.0035
	1555+001	DW 1555+00	1976.86	14	22	34	15	57	51.43418	0.00033	-0	1	50.4203	0.0052
	1611+343	DA 406	1977.37	16	30	44	16	13	41.06409	0.00025	34	12	47.9082	0.0026
	1633+382	GC 1633+38	1980.04	10	32	32	16	35	15.49283	0.00013	38	8	4.4985	0.0013
	1638+398	NRAO 512	1979.50	16	54	53	16	40	29.63258	0.00015	39	46	46.0278	0.0014
	1641+399	3C 345	1978.28	19	54	72	16	42	58.80983	0.00013	39	48	36.9928	0.0012
	1656+053	DW 1656+05	1978.85	9	12	16	16	58	33.44733	0.00061	5	15	16.4383	0.0093
	1717+178	GC 1717+17	1980.07	8	19	19	17	19	13.04837	0.00022	17	45	6.4352	0.0037
	1730-130	NRAO 530	1979.13	14	28	28	17	33	2.70553	0.00018	-13	4	49.5460	0.0038
	1738+476	OT 465	1978.70	4	6	6	17	39	57.12566	0.00078	47	37	58.3768	0.0045
	1741-038	P 1741-038	1977.44	13	25	35	17	43	58.85676	0.00027	-3	50	4.6252	0.0046
	1749+701	1749+701	1979.45	11	16	16	17	48	32.83875	0.00060	70	5	50.7750	0.0032
	1807+698	3C 371	1978.96	18	48	57	18	6	50.67971	0.00026	69	49	28.1088	0.0010
	1821+107	P 1821+10	1980.06	7	14	14	18	24	2.85524	0.00013	10	44	23.7698	0.0048
S	1901+319	3C 395	1979.61	4	5	5	19	2	55.83921	0.08749	31	59	40.3952	0.7978
	1921-293	OV-236	1978.47	4	11	9	19	24	51.05564	0.00044	-29	14	30.1133	0.0055
	1933-400	P 1933-400	1979.86	4	7	7	19	37	16.21675	0.00059	-39	58	1.5528	0.0061
	1958-179	OV-198	1979.30	8	18	17	20	0	57.09073	0.00037	-17	48	57.6761	0.0055
	2021+614	OW 637	1978.47	11	18	19	20	22	6.68158	0.00095	61	36	58.8193	0.0055
	2029+547	OW 551	1980.02	3	8	8	20	31	47.95842	0.00040	54	55	3.1495	0.0037
	2030+121	P 2029+121	1980.06	9	14	14	20	31	54.99410	0.00019	12	19	41.3436	0.0034
S	2037+511	3C 418	1979.64	3	5	5	20	38	36.94366	0.06543	51	19	11.8561	0.4130
	2113+293	B2 2113+29B	1979.83	11	19	19	21	15	29.41343	0.00015	29	33	38.3663	0.0024
	2134+004	P 2134+004	1977.17	14	26	40	21	36	38.58616	0.00017	0	41	54.2150	0.0038
	2145+067	P 2145+06	1978.52	21	31	39	21	48	5.45853	0.00011	6	57	38.6057	0.0023
	2149+056	OX 082	1980.06	8	16	16	21	51	37.87530	0.00017	5	52	12.9556	0.0034
	2155-152	OX-192	1978.66	4	6	6	21	58	6.28154	0.00072	-15	1	9.3263	0.0109
	2200+420	VRO 42.22.01	1978.05	27	46	60	22	2	43.29125	0.00016	42	16	39.9839	0.0014
	2230+114	CTA 102	1978.45	10	13	13	22	32	36.40897	0.00050	11	43	50.9052	0.0068
	2234+282	GC 2234+28	1980.04	10	23	22	22	36	22.47076	0.00013	28	28	57.4168	0.0018
	2243-123	OY-172.6	1979.27	14	22	22	22	46	18.23184	0.00014	-12	6	51.2764	0.0033
	2245-328	P 2245-328	1979.55	7	20	20	22	48	38.68551	0.00024	-32	35	52.1861	0.0036
	2251+158	3C 454.3	1977.56	17	37	50	22	53	57.74779	0.00015	16	8	53.5658	0.0027
	2253+417	GC 2253+41	1980.11	3	11	11	22	55	36.70799	0.00019	42	2	52.5370	0.0030
	2320-035	P 2320-035	1980.00	5	8	7	23	23	31.95363	0.00021	-3	17	5.0216	0.0044
	2345-167	P 2345-16	1977.98	16	20	24	23	48	2.60846	0.00035	-16	31	12.0233	0.0052
S	2352+495	DA 611	1979.66	3	5	5	23	55	9.45244	0.07304	49	50	8.4831	0.2230

* In barycentric coordinates.
S Short-baseline data only.
R Inadequate redundancy fewer than 5 delay observations.

Table 4. Root-mean-square residuals for 1971–1980 observing sessions in the standard fit

Date	Antennas	RMS residuals Delay (ns)	RMS residuals Delay rate (psec/sec)
71/ 8/28	14-62	-	0.27
71/ 9/ 1	14-62	-	0.16
71/ 9/ 6	14-62	-	0.26
71/ 9/10	14-62	-	0.33
73/ 4/30	14-62	-	0.26
73/ 9/ 8	14-62	-	0.45
74/ 2/15	14-62	-	0.26
74/ 4/21	14-62	-	0.33
74/ 6/21	14-62	-	0.43
74/ 8/ 6	14-62	-	0.28
77/ 1/12	11-43	0.52	0.74
77/ 1/21	11-14-43	0.62	0.36
77/ 1/31	14-63	0.87	0.34
77/ 2/ 1	11-43	0.75	0.56
77/ 2/13	11-14-43	0.84	0.33
77/ 2/28	11-14-43	0.77	0.32
77/ 4/13	14-63	0.97	0.34
78/ 1/14	14-43	0.29	0.17
78/ 1/24	14-43	0.35	0.37
78/ 2/12	14-43	0.18	0.63
78/ 2/24	OV-HY	0.29	0.05
78/ 5/15	14-43	0.48	0.66
78/ 7/30	14-63	0.64	0.45
78/ 9/ 3	14-43	0.44	0.50
78/ 9/ 4	14-63	1.03	0.41
78/10/27	14-43	0.46	0.14
78/10/30	14-63	0.49	0.83
78/11/ 4	14-43	0.25	0.10
78/11/ 5	14-63	0.73	0.85
78/12/31	14-43	0.34	0.14
79/ 6/ 6	13-14	0.22	0.16
79/ 7/21	13-14	0.17	0.11
79/ 8/26	13-14	0.10	0.09
79/ 9/18	13-14	0.14	0.06
79/11/15	62-63	0.09	0.19
79/11/23	14-43	0.53	0.20
79/11/25	14-63	0.57	0.13
79/12/20	14-43	0.50	0.14
79/12/21	14-63	0.56	0.11
79/12/27	14-63	0.46	0.13
79/12/29	14-43	0.49	0.11
80/ 1/12	14-43	0.44	0.21
80/ 1/25	14-63	0.48	0.08
80/ 1/27	14-43	0.43	0.15
80/ 2/13	14-63	0.50	0.10
80/ 2/14	14-43	0.32	0.12
80/ 2/23	14-43	0.42	0.14
80/ 2/24	14-63	0.40	0.11
Average		0.52	0.30

Table 5. Source position repeatability results

Source	Average epoch		* Difference (mas)		RSS errors (mas)		Diff/RSS	
	Part 1	Part 2	RA	Dec	RA	Dec	RA	Dec
P 0104-408	78.74	80.06	-1.6	2.2	4.8	5.5	-0.33	0.40
P 0106+01	77.24	79.93	1.8	-0.4	4.7	5.4	0.38	-0.07
DA 55	77.74	79.92	3.2	3.7	3.9	3.0	0.83	1.23
DW 0224+67	76.34	79.84	0.6	-1.2	3.6	4.1	0.18	-0.29
NRAO 140	75.86	79.94	-1.8	3.9	4.4	3.7	-0.40	1.05
DW 0742+10	77.73	80.00	5.6	-6.7	4.0	4.3	1.40	-1.55
4C 39.25	77.85	80.02	2.2	0.3	2.6	2.3	0.84	0.13
P 1127-14	77.98	80.01	1.7	0.4	4.6	5.7	0.37	0.07
P 1144-379	78.12	80.00	0.0	-0.2	4.5	5.2	0.00	-0.04
OR 103	77.95	80.00	1.8	1.4	7.3	7.6	0.24	0.18
NRAO 512	77.83	80.04	-4.2	2.7	5.2	5.0	-0.80	0.54
3C 345	76.06	80.01	-1.5	0.8	3.3	3.5	-0.46	0.23
NRAO 530	78.31	79.95	-1.9	12.1	7.1	7.9	-0.27	1.54
3C 371	77.09	79.98	-6.6	2.2	4.9	4.1	-1.34	0.54
P 2145+06	76.62	80.03	-0.9	2.4	4.8	5.3	-0.19	0.45
VRO 42.22.01	76.82	79.92	-0.6	3.9	3.4	3.0	-0.16	1.31
OY-172.6	78.64	80.02	3.5	-4.0	5.0	6.0	0.71	-0.66
		RMS =	2.9	4.0		χ^2/N =	0.43	0.63

* Difference : 1971-78 position minus 1979-80 position.

Table 6. Station locations from 1971–1980 VLBI data

Station	CIO cylindrical coordinates		
	Equatorial (m)	Longitude (degrees)	Polar (m)
DSS11	5206339.890±0.562	243.1505770±0.0000055	3673763.464±0.322
DSS13	5215483.938 0.022	243.2051172 0.0000001	3660956.516 0.019
* DSS14	5203996.766 -	243.1104671 -	3677052. -
DSS43	5205251.060 0.087	148.9812611 0.0000010	-3674749.217 0.141
DSS62	4860817.898 0.162	355.6321691 0.0000028	4116905.731 0.250
DSS63	4862451.030 0.158	355.7519856 0.0000028	4115108.644 0.249
OVRO	5085449.442 0.093	241.7172850 0.0000010	3838603.826 0.092
HAYST	4700479.623 0.109	288.5118354 0.0000012	4296882.190 0.115

* Reference station.

Table 7. Selected baselines from 1971–1980 VLBI data

Baseline	CIO rectangular coordinates (m)			Length (m)
	x	y	z	
DSS14-DSS43	-2107273.066±0.117	7323703.098±0.042	-7351801.217±0.141	10588966.331±0.101
DSS14-DSS63	7202713.569 0.141	4281160.656 0.248	438056.644 0.249	8390429.843 0.043
OVRO-HAYST	3902005.520 0.027	21082.369 0.064	458278.364 0.069	3928881.596 0.025
DSS14-DSS11	2191.480 0.628	-3737.042 0.412	-3288.537 0.322	5438.983 0.301
DSS14-DSS13	2492.103 0.014	-14135.491 0.020	-16095.484 0.019	21565.883 0.013
DSS62-DSS63	2392.199 0.034	10015.216 0.013	-1797.088 0.023	10452.593 0.016

Table 8. Baseline formal uncertainties* from fit to 1971–1980 VLBI data

Baseline	Formal uncertainty (cm) x	y	z	Length
DSS11-DSS13	62.8	41.2	32.3	32.2
DSS14	62.8	41.2	32.2	30.1
DSS43	63.9	41.2	34.9	38.0
DSS62	64.4	48.2	40.8	66.8
DSS63	64.3	48.1	40.7	66.7
OVRO	63.5	42.2	33.5	30.7
HAYST	63.5	42.7	34.2	63.9
DSS13-DSS14	1.4	2.0	1.9	1.3
DSS43	11.8	4.7	14.2	10.3
DSS62	14.5	24.9	25.1	5.6
DSS63	14.1	24.9	25.0	4.7
OVRO	(9.0	9.0	9.0	9.0)
HAYST	9.4	11.1	11.3	9.3
DSS14-DSS43	11.7	4.2	14.1	10.1
DSS62	14.5	24.8	25.0	5.3
DSS63	14.1	24.8	24.9	4.3
OVRO	9.1	9.2	9.2	9.1
HAYST	9.5	11.2	11.5	9.4
DSS43-DSS62	16.4	25.6	31.4	32.1
DSS63	16.1	25.6	31.3	31.9
OVRO	14.9	10.1	16.8	13.8
HAYST	14.9	12.5	19.8	15.4
DSS62-DSS63	3.4	1.3	2.3	1.6
OVRO	17.1	26.5	26.7	10.6
HAYST	17.2	26.0	26.0	15.5
DSS63-OVRO	16.8	26.4	26.6	10.2
HAYST	16.9	26.0	25.9	15.4
OVRO-HAYST	2.7	6.4	6.9	2.5

* In CIO rectangular coordinates.

Table 9. Polar motion and UT1 values* from 1971–1980 VLBI data

Date	Mean Julian date	Polar motion, mas x		Polar motion, mas y		UT1-UTC, msec	
1971/ 8/29	41192.17					-76.7	± 1.3
1971/ 9/ 2	41196.19					-76.0	1.4
1971/ 9/ 7	41201.14					-82.1	1.2
1971/ 9/11	41205.08					-82.6	1.6
1973/ 4/30	41802.87					405.0	2.0
1973/ 9/ 8	41933.85					54.2	3.1
1974/ 2/16	42094.06					577.6	1.9
1974/ 4/21	42158.83					383.4	2.4
1974/ 6/21	42219.58					215.7	3.6
1974/ 8/ 6	42265.87					113.2	4.1
1977/ 1/12	43155.59			100.7	± 8.8	632.3	2.4
1977/ 1/21	43164.43			122.8	6.5	605.6	1.4
1977/ 1/31	43174.34					577.9	0.9
1977/ 1/31	43174.84	-171.0	±15.2	151.3	8.3		
1977/ 2/ 1	43175.44					577.8	2.4
1977/ 2/13	43187.31			190.4	5.9	544.4	1.2
1977/ 2/28	43202.19			240.6	3.1	500.3	0.8
1977/ 4/13	43246.23	-184.6	14.1			362.3	0.8
1978/ 1/14	43522.67			21.5	3.0	602.0	0.8
1978/ 1/24	43532.16			38.0	3.1	576.2	0.8
1978/ 2/12	43551.56			72.7	3.0	511.1	0.7
1978/ 2/25	43564.92	-177.2	...	114.6	...	467.2	...
1978/ 5/15	43643.53			389.0	3.2	203.1	0.8
1978/ 7/31	43720.16	72.2	15.8			25.5	0.9
1978/ 9/ 3	43754.89					-53.5	1.8
1978/ 9/ 4	43755.42	166.9	25.4	422.0	7.4		
1978/ 9/ 4	43755.70					-53.8	1.1
1978/10/28	43809.02					-205.5	0.7
1978/10/28	43809.58	240.8	9.1	256.9	2.7		
1978/10/30	43811.20					-212.0	0.7
1978/11/ 4	43816.76					-229.1	0.6
1978/11/ 5	43817.84	241.5	7.6	223.2	2.3		
1978/11/ 6	43818.18					-233.8	0.5
1978/12/31	43873.44			86.8	2.9	-400.7	0.7
1979/11/23	44200.80					-256.6	0.6
1979/11/25	44202.64	130.9	5.9	320.5	2.3		
1979/11/26	44203.30					-261.7	0.5
1979/12/20	44227.56					-325.7	...
1979/12/20	44227.78	139.4	...	273.1	...		
1979/12/21	44228.71					-328.2	...
1979/12/27	44234.79					-343.8	0.5
1979/12/28	44235.54	134.3	6.5	259.7	2.0		
1979/12/29	44236.66					-349.3	0.5
1980/ 1/12	44250.83			237.3	2.9	616.0	0.7
1980/ 1/26	44264.00					581.0	0.5
1980/ 1/26	44264.75	110.2	7.1	214.1	2.2		
1980/ 1/27	44265.62					576.6	0.6
1980/ 2/14	44283.21					539.8	0.5
1980/ 2/14	44283.40	73.1	5.7	191.2	1.7		
1980/ 2/14	44283.92					537.4	0.4
1980/ 2/23	44292.85					510.9	0.5
1980/ 2/24	44293.04	58.3	6.4	184.5	1.9		
1980/ 2/24	44293.54					510.3	0.5

* These UT values include short-period tidal effects.

Table 10. Zenith tropospheric delays from 1971–1980 VLBI data

Station	Date	Mean Julian date	Zenith delay (cm)
DSS11	77/ 1/12	43155.59	207.3 ± 5.3
	77/ 1/21	43164.48	212.6 4.3
	77/ 2/ 1	43175.44	210.3 4.9
	77/ 2/13	43187.33	216.1 4.1
	77/ 2/28	43202.11	211.1 5.9
DSS14	71/ 8/29	41192.17	222.0 ± 4.0
	71/ 9/ 2	41196.19	214.0 3.9
	71/ 9/ 7	41201.14	214.2 4.0
	71/ 9/11	41205.08	213.3 4.2
	73/ 4/30	41802.87	209.8 5.5
	73/ 9/ 8	41933.85	211.7 5.5
	74/ 2/16	42094.06	211.2 5.5
	74/ 4/21	42158.83	206.8 4.8
	74/ 6/21	42219.58	211.8 5.9
	74/ 8/ 6	42265.87	216.3 5.9
	77/ 1/21	43164.39	213.1 5.0
	77/ 1/31	43174.34	200.6 2.8
	77/ 2/13	43187.30	215.9 4.1
	77/ 2/28	43202.20	208.9 1.6
	77/ 4/13	43246.23	209.0 1.8
	78/ 1/14	43522.67	215.8 1.2
	78/ 1/24	43532.16	209.5 1.7
	78/ 2/12	43551.56	208.5 0.6
	78/ 5/15	43643.53	210.9 1.7
	78/ 7/31	43720.16	206.2 2.0
	78/ 9/ 3	43754.89	215.2 2.1
	78/ 9/ 4	43755.70	211.9 4.7
	78/10/28	43809.02	209.0 0.7
	78/10/30	43811.20	208.2 2.3
	78/11/ 4	43816.76	211.4 0.8
	78/11/ 5	43817.90	210.7 2.5
	78/11/ 6	43818.38	210.7 3.6
	78/12/31	43873.23	208.5 0.9
		43873.67	206.7 0.8
	79/11/23	44200.70	210.9 3.1
		44200.87	210.0 1.8
	79/11/26	44203.13	209.1 1.1
		44203.62	209.0 1.8
	79/12/20	44227.23	212.0 1.3
		44227.71	207.7 0.8
	79/12/21	44228.71	210.9 2.7
	79/12/27	44234.79	207.4 2.9
	79/12/29	44236.66	210.7 1.0
	80/ 1/12	44250.69	219.7 1.0
	80/ 1/13	44251.12	217.4 1.4
	80/ 1/25	44263.73	212.1 3.3
	80/ 1/26	44264.09	206.6 2.1
	80/ 1/27	44265.54	205.6 1.8
		44265.69	204.1 1.0
	80/ 2/13	44282.80	211.5 1.5
	80/ 2/14	44283.09	210.7 0.9
		44283.50	211.6 1.2
		44283.92	215.9 1.3
	80/ 2/23	44292.85	210.1 0.8
	80/ 2/24	44293.54	208.9 1.7

Table 10 (contd)

Station	Date	Mean Julian date	Zenith delay (cm)
DSS43	77/ 1/12	43155.59	226.5 ± 4.9
	77/ 1/21	43164.43	219.0 2.9
	77/ 2/ 1	43175.44	227.8 4.1
	77/ 2/13	43187.31	226.7 2.7
	77/ 2/28	43202.19	226.2 1.5
	78/ 1/14	43522.67	223.2 1.3
	78/ 1/24	43532.16	229.1 2.3
	78/ 2/12	43551.56	222.4 0.6
	78/ 5/15	43643.53	224.0 2.4
	78/ 9/ 3	43754.89	227.6 2.0
	78/10/28	43809.02	227.2 0.8
	78/11/ 4	43816.76	220.4 0.8
	78/12/31	43873.23	225.6 0.8
		43873.67	224.6 0.7
	79/11/23	44200.80	234.0 1.9
	79/12/20	44227.23	221.6 0.9
		44227.56	221.0 0.8
		44227.88	221.9 1.2
	79/12/29	44236.56	230.7 1.0
		44236.76	228.6 2.0
	80/ 1/12	44250.83	221.0 0.8
	80/ 1/27	44265.62	223.5 1.1
	80/ 2/14	44283.92	218.7 0.9
	80/ 2/23	44292.85	223.9 0.8
DSS62	71/ 8/29	41192.17	220.0 ± 4.4
	71/ 9/ 2	41196.19	227.0 4.3
	71/ 9/ 7	41201.14	225.2 4.6
	71/ 9/11	41205.08	231.0 4.9
	73/ 4/30	41802.87	223.0 5.5
	73/ 9/ 8	41933.85	224.7 5.5
	74/ 2/16	42094.06	218.6 4.7
	74/ 4/21	42158.83	223.4 5.5
	74/ 6/21	42219.58	224.3 5.5
	74/ 8/ 6	42265.87	230.1 5.7
DSS63	77/ 1/31	43174.34	215.7 ± 4.2
	77/ 4/13	43246.23	215.7 2.4
	78/ 7/31	43720.16	219.0 3.0
	78/ 9/ 4	43755.70	230.4 4.9
	78/10/30	43811.20	217.3 3.0
	78/11/ 5	43817.90	220.1 3.0
	78/11/ 6	43818.38	221.1 2.7
	79/11/26	44203.30	217.0 1.1
	79/12/21	44228.64	211.0 1.7
		44228.82	215.2 2.4
	79/12/27	44234.79	219.7 1.4
	80/ 1/26	44264.00	221.2 1.7
	80/ 2/13	44282.80	216.9 1.6
	80/ 2/14	44283.09	218.2 1.1
		44283.50	219.8 1.3
	80/ 2/24	44293.54	222.6 1.4
OVRO	78/ 2/25	43564.17	203.6 ± 1.7
		43564.67	205.0 1.2
	78/ 2/26	43565.03	196.6 3.1
		43565.50	202.0 1.7
HAYST	78/ 2/25	43564.17	228.4 ± 1.1
		43564.67	227.3 0.7
	78/ 2/26	43565.03	228.9 1.4
		43565.50	228.7 0.9

Table 11. Differential clock parameters* from 1971–1980 VLBI data

Station	Date	Epoch(ns)	Rate(psec/sec)	dR/dt(1/sec) *10**18
DSS11	77/ 2/13 7:49	-8604± 43	0.462±1.30	...
	77/ 2/28 2:38	-12740 19	-1.398 0.565	...
DSS13	79/ 6/ 6 7:15	-249±0.3	-1.091±0.020	...
	79/ 7/21 13:01	-219 0.1	-2.364 0.016	2.3 ± 1.6
	79/ 8/26 5:30	758 0.3	-1.126 0.013	...
	8:03	758 0.7	-1.075 0.048	...
	8:51	756 1.1	-1.283 0.098	...
	79/ 9/18 5:27	1620 0.2	-2.134 0.006	...
DSS14	77/ 1/21 9:15	1371± 3	0.249±0.280	...
	77/ 2/13 6:43	13 47	-0.613 1.65	...
	9:33	-41 43	-0.633 1.32	...
DSS43	77/ 1/12 14:06	-1479± 4	-3.295±0.240	...
	77/ 1/21 9:30	12170 2	-4.406 0.252	...
	12:49	12223 3	-4.596 0.257	...
	77/ 2/ 1 9:18	28444 4	-2.160 0.198	...
	12:02	28555 3	-1.978 0.328	...
	77/ 2/13 2:43	29245 22	-2.200 0.656	...
	6:46	29272 46	-3.207 1.49	...
	9:46	29275 43	-2.888 1.31	...
	77/ 2/28 3:18	29961 1	-2.334 0.019	...
	9:00	30047 1	-2.217 0.057	...
	78/ 1/14 8:47	24356 1	1.333 0.048	...
	16:03	24324 0.8	1.319 0.020	...
	19:31	24326 1	1.219 0.019	...
	78/ 1/24 3:49	19393 1	7.014 0.014	...
	78/ 2/12 13:21	15482 0.6	-1.054 0.008	-5.6 ± 0.5
	78/ 5/15 9:16	10130 1	0.989 0.039	...
	12:16	10128 1	0.760 0.029	...
	16:22	10125 1	1.243 0.028	...
	78/ 9/ 3 21:20	4539 1	-3.297 0.018	...
	78/10/27 19:26	8786 0.6	...	...
	78/10/28 0:23	...	2.018 0.006	...
	5:20	8786 0.6	...	...
	78/11/ 4 18:11	10306 0.7	2.028 0.012	...
	78/12/31 8:28	16003 0.6	0.211 0.004	...
	22:29	16010 1	0.126 0.028	...
	79/11/23 16:43	38096 2	-0.467 0.048	...
	19:50	38100 2	-0.285 0.147	...
	22:16	38138 1	-0.318 0.057	...
	79/12/20 13:22	36494 0.6	-0.577 0.003	...
	79/12/29 13:31	35896 0.7	-1.057 0.017	...
	18:08	35896 0.8	-1.061 0.019	...
	80/ 1/12 13:18	35904 0.7	0.305 0.009	...
	23:13	35905 0.6	0.357 0.006	...
	80/ 1/27 14:54	36511 0.8	0.810 0.014	...
	80/ 2/14 20:03	33524 0.9	-0.151 0.018	...
	80/ 2/15 0:33	33525 0.8	-0.131 0.015	...
	80/ 2/23 20:21	33349 0.6	-0.199 0.002	...

Table 11 (contd)

Station	Date	Epoch(ns)	Rate(psec/sec)	dR/dt(1/sec) *10**18
DSS62	71/ 8/29 4:11	...	-1.625±0.056	...
	71/ 9/ 2 4:29	...	-1.741 0.049	...
	71/ 9/ 7 3:16	...	-2.086 0.052	...
	71/ 9/10 22:46	...	-1.959 0.080	...
	71/ 9/11 5:44	...	-1.529 0.086	...
	73/ 4/30 17:52	...	-6.532 0.138	...
	22:03	...	-6.471 0.085	...
	73/ 9/ 8 20:25	...	6.622 0.138	...
	74/ 2/16 1:27	...	7.711 0.082	...
	74/ 4/21 19:52	...	6.453 0.096	...
	74/ 6/21 14:02	...	-12.113 0.143	...
	74/ 8/ 6 20:50	...	-0.951 0.176	...
DSS63	77/ 1/31 6:34	-23453± 3	4.869±0.143	...
	9:12	-23440 4	5.561 0.388	...
	10:52	-23562 7	4.420 1.58	...
	77/ 4/13 3:03	-41060 6	-2.520 0.180	...
	6:45	-41068 3	-2.752 0.248	3.1 ± 6.3
	78/ 7/31 0:34	-6499 2	-0.020 0.135	...
	3:44	...	0.020 0.105	...
	5:10	-6501 2	...	...
	6:46	...	0.047 0.072	...
	78/ 9/ 4 13:36	...	-1.751 0.221	...
	16:54	-7667 4	...	...
	18:15	...	-1.764 0.134	...
	78/10/30 2:06	-3799 2	0.115 0.035	...
	6:35	-3798 2	0.129 0.031	...
	78/11/ 5 22:28	-3665 1	0.384 0.013	...
	78/11/ 6 9:37	-3662 1	0.312 0.017	...
	79/11/15 22:44	7019 0.1	1.753 0.006	...
	79/11/25 21:46	8043 0.6	0.423 0.015	...
	79/11/26 4:45	8050 0.6	0.354 0.014	...
	12:03	8051 0.8	0.315 0.013	...
	17:01	8049 3.7	0.353 0.061	...
	79/12/21 16:58	8745 1	0.418 0.012	...
	79/12/27 13:10	9302 0.9	0.545 0.007	...
	21:03	9302 0.8	0.589 0.023	...
	79/12/28 2:15	9303 0.8	0.561 0.019	...
	80/ 1/25 21:44	10275 0.9	0.595 0.011	...
	80/ 1/26 9:36	10256 18	0.988 0.391	...
	80/ 2/13 22:17	7649 0.5	-0.098 0.009	...
	80/ 2/14 10:01	7650 0.6	-0.146 0.005	...
	80/ 2/24 13:00	7758 0.8	0.086 0.007	...
OVRO	78/ 2/25 3:36	-11424±0.3	-0.109±0.006	...
	17:28	-11424 0.3	-0.180 0.004	...
	78/ 2/26 5:38	-11428 0.7	-0.092 0.010	...
	15:53	-11420 0.5	-0.188 0.005	...

* All sessions are relative to the DSS14 clock except for the OVRO/HAYST session (relative to HAYST) and the 77/1/21, 77/2/13 11/14/43 sessions (relative to DSS11 for some portions).

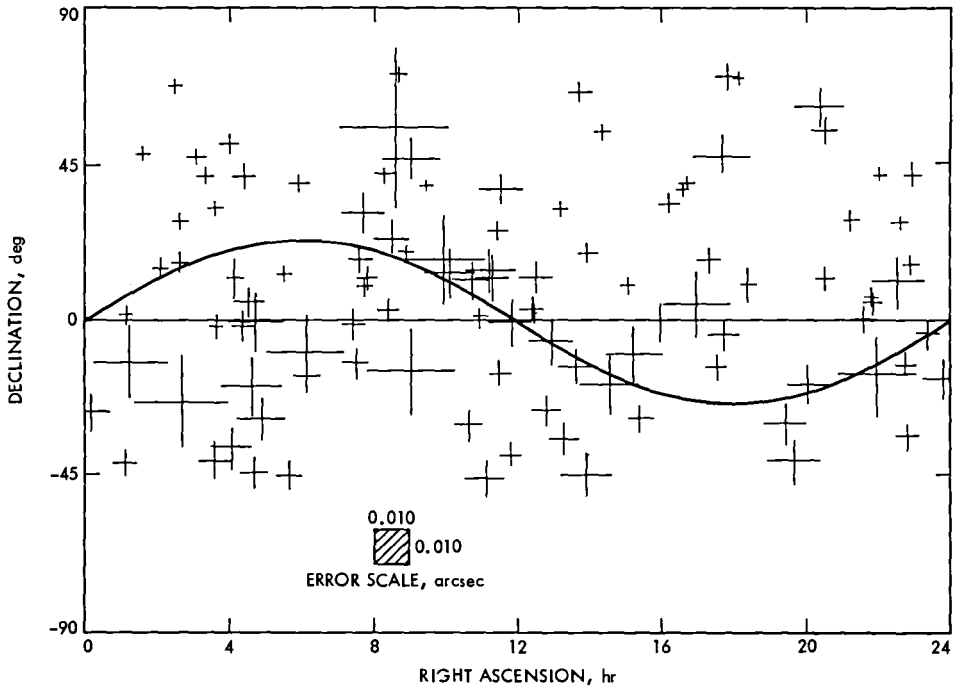


Fig. 1. Source positions and their formal errors from a fit to 1971–80 VLBI data

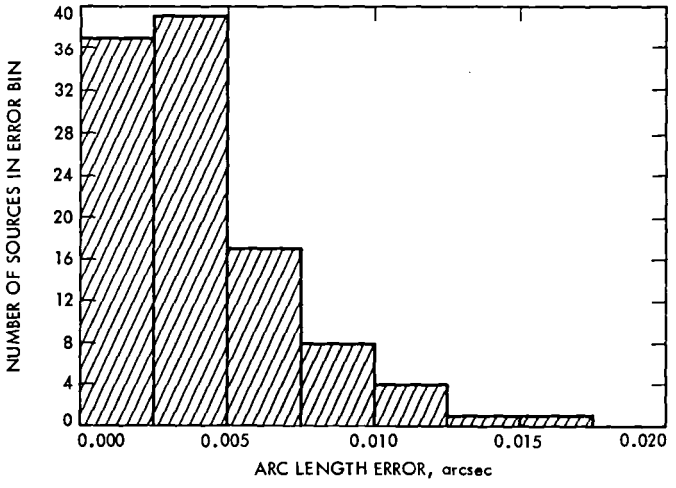


Fig. 2. Histogram of right ascension formal errors in the fit to 1971–80 data

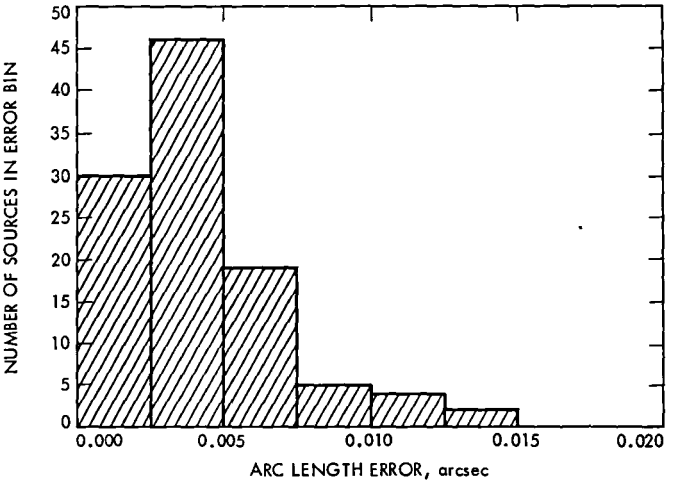


Fig. 3. Histogram of declination formal errors in the fit to 1971–80 data

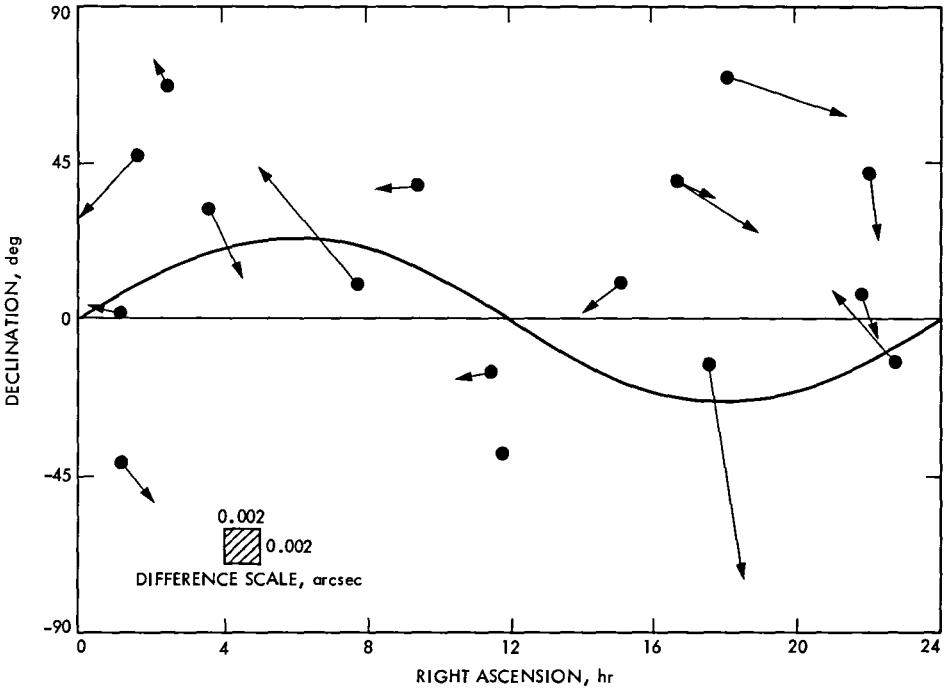


Fig. 4. Source position differences: 1971–78 minus 1979–80 for 17 sources. The difference vector points away from the filled circle specifying the 1979–80 source position. Note the scale change for position and difference.

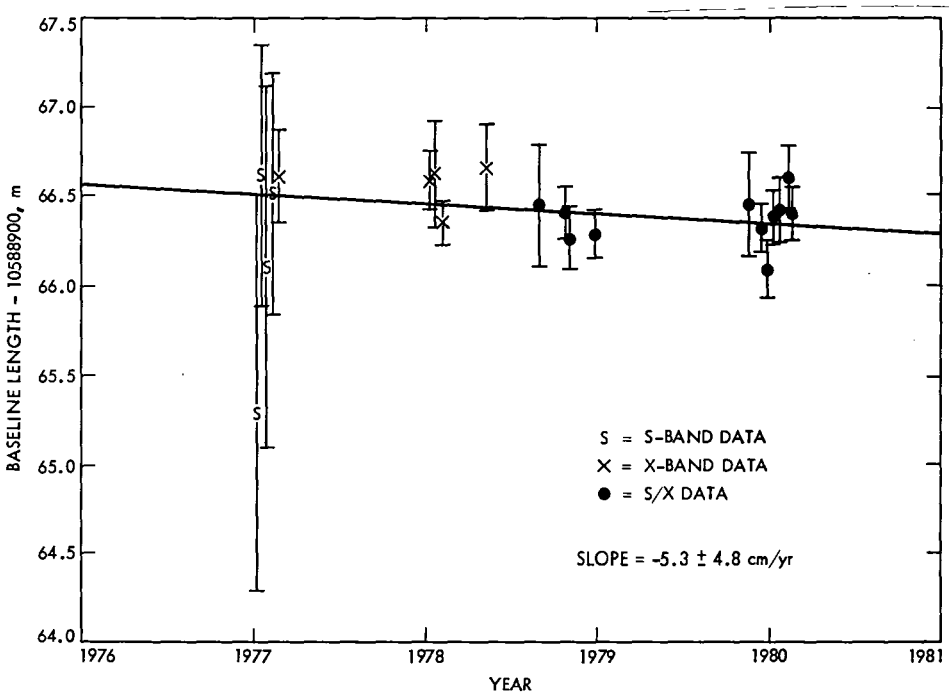


Fig. 5. Length of the California-Australia baseline as a function of time from a fit to 1977–80 VLBI data

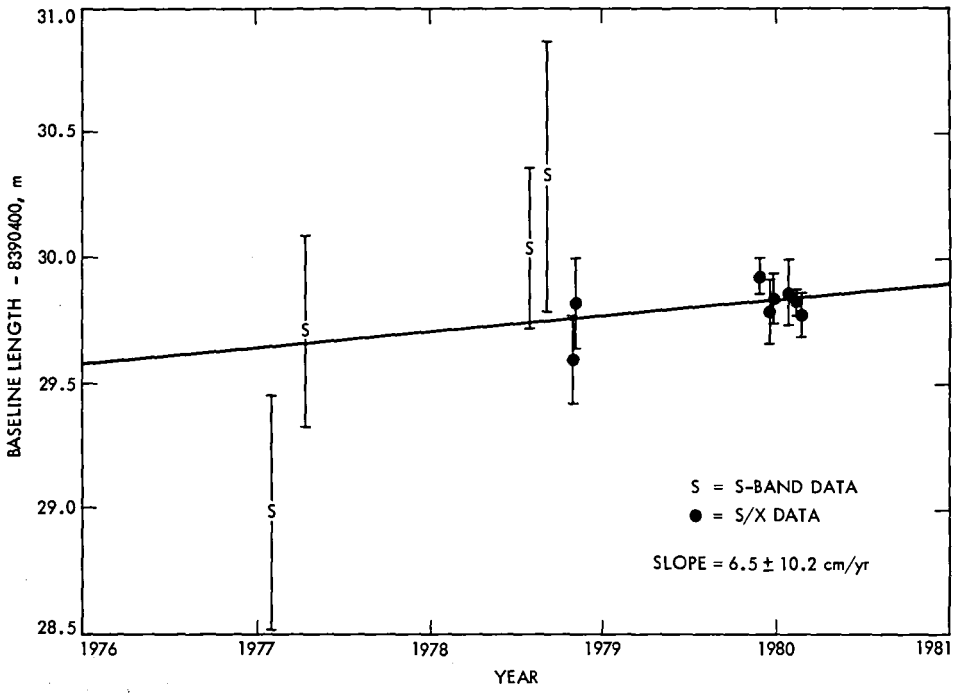


Fig. 6. Length of the California-Spain baseline as a function of time from a fit to 1977–80 VLBI data

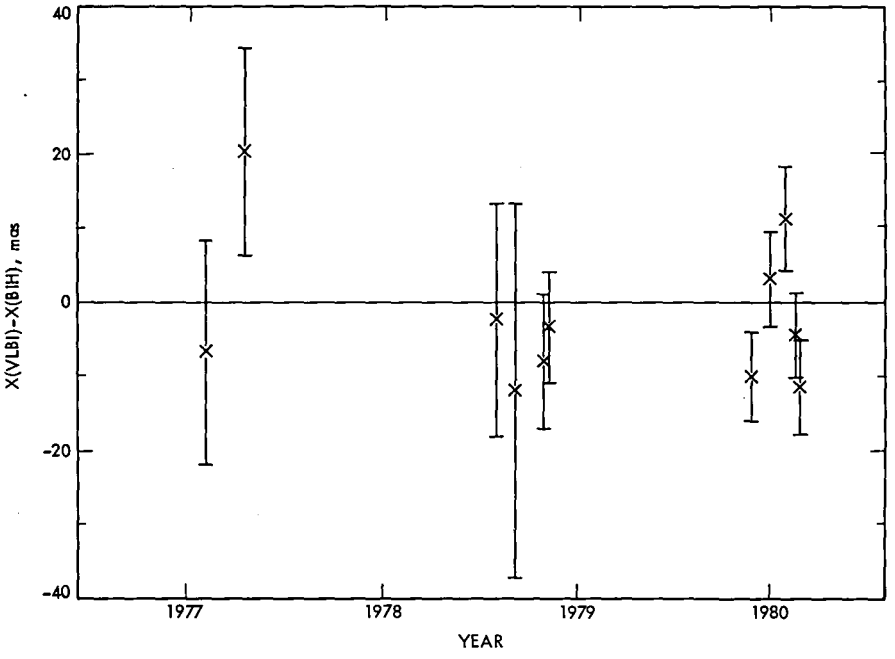


Fig. 7. Polar motion X-component results from 1977–1980 VLBI data

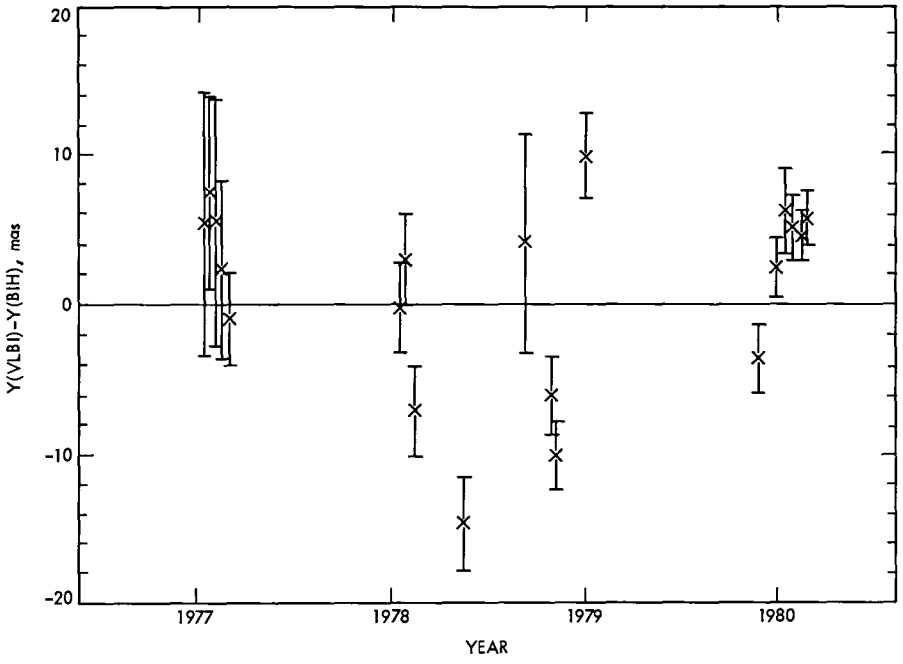


Fig. 8. Polar motion Y-component results from 1977–1980 VLBI data

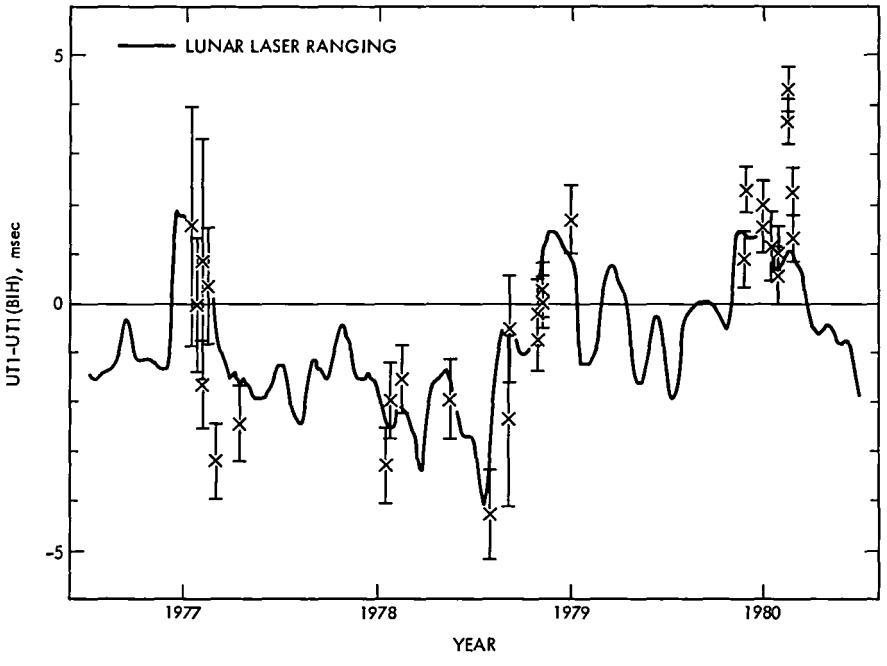


Fig. 9. UT1 results from 1977–1980 VLBI data

Limits to Arraying

J. W. Layland
TDA Planning Office

Arraying has been used, tested, analyzed, and generally poked about for the past decade or so. We have used one technique, or block diagram, for arraying for telemetry in the DSN, analyzed a few others, and seen still others used in radio astronomy applications. Looking to the future, there are a number of applications where our present technique will not work as well as we would like, and other techniques must be tried. This article develops (approximately) the feasibility limit to various possible array configurations. It is not intended as a "last-word" analysis, but rather as a guide to which areas would be the most fruitful for future analysis and development.

I. Introduction

Arraying has been used, tested, analyzed, and generally poked about for the past decade or so. We have in hand a large collection of answers to questions we have asked, but do not seem to have a generalized description of where arraying can be applied, and where it cannot. Can we expand an array to an arbitrarily large number of arbitrarily small antennas? The answer, of course, is "No," and the actual limits encountered depend upon the complexity of the signal processing/combining employed.

Arraying of several antennas to combine their effective receiving aperture requires that the narrowband spacecraft signal as it is received at each antenna be precisely aligned in delay and phase with the corresponding signal as received at each of the other antennas so that they can be coherently added. This alignment can be mechanized in a variety of ways. The mechanization used in the DSN today is illustrated in Fig. 1, where a phase locked loop is associated with each individual antenna to determine the phase of the received carrier component of the signal, and thus provide a coherent reference for translating the received carrier to DC (zero-phase). This reference is also used for coherently detecting the modulation, which consists of subcarrier times data. With the carrier phase thus eliminated from further concern, the baseband signal is filtered to a bandwidth consistent with reasonably precise representation of the modulated square wave subcarrier, sampled, delay-adjusted into agreement with the corresponding signals from the other antennas, and weighted by its respective strength prior to being added to the other signals. The combined result is then subjected to subcarrier and symbol detection.

The current DSN arraying is subject to a variety of losses, including that from noise in the carrier reference used to detect the signals at each of the antennas. Two methods of decreasing the noise in the carrier reference signals have been proposed and studied – these are carrier arraying and data-aided carrier tracking. In carrier arraying, the remnant carrier signal is isolated from the modulation sidebands, and the

phase of the received carrier in each of the antennas is tracked relative to that in one reference antenna (or possibly to a combined carrier signal), and this relative phase measurement used to align the carrier signals from each antenna so that they can be combined directly. The combined carrier signal is stronger than the signal at each individual antenna, and is thus more easily tracked. Tracking of the relative phase between signals as handled at each of the antennas is also easier because it must follow differential doppler only, at a bandwidth which is on the order of 10^{-3} times the full topocentric doppler. Data-aided carrier tracking improves the ability to develop a carrier reference for each individual antenna by reconstructing a stronger carrier-equivalent signal from the data sidebands.

The Very Large Array (VLA) is a radio astronomy observatory near Socorro, New Mexico, which consists of twenty-seven 25-meter antennas (plus a spare). Its primary product is interferometric maps of various radio-bright objects, but it also happens to be able to produce a combined signal output by a mechanism which is crudely diagrammed in Fig. 2. This is "total-spectrum combining," where both the differential delay and differential phase of the individual antennas are compensated for in one combining step, rather than in two as in other mechanisms.

One additional mechanism which deserves consideration here incorporates the subcarrier demodulation and symbol stream detection into the processing which is done on individual antennas before the signals are transferred to a central point for combining. This adds somewhat to the complexity of the equipment on individual antennas, but makes the final combining process extremely simple – a matter of aligning, weighting, and adding the discrete-time symbol stream. Symbol detection on individual antennas also facilitates the measurement of the signal strength as received at each antenna, which is needed for weighting the combined inputs, and may also facilitate monitoring the health of the combining process.

A word of warning to the reader is in order here – the analysis which follows is at an architectural conceptual level, fraught with approximations and use of nonrealizeable square passband filters, etc. Exact implementation-specific analyses have been done elsewhere for some, but not all, of the configurations discussed. Choice of filters, etc., can easily make the exact results differ by a dB (or a little more) from the approximate results to be presented here.

II. Arraying Mechanisms and Limiting Conditions

The ensemble of conditions pertinent to arraying is large, and hence some arbitrariness will limit the set of conditions first considered to symmetric arrays only, such as the VLA. Nonsymmetric arrays will be briefly discussed later. The performance boundaries are a function of two parameters – the data rate and the individual antenna size relative to that of the total array. We assume that the overall signal into the array is just adequate to support the desired data rate. Thus, as the data rate decreases, the signal power as received by either the total array or the individual antennas decreases, and the task of synchronizing the signals from them becomes increasingly difficult. Similarly, as the size of the individual antennas decreases, the signal power in each of them decreases correspondingly, and the task of arraying them likewise becomes more difficult. We will *assume*, for a concrete example, that adequate synchronizing of the signals being arrayed requires a loop margin of +15 dB in the phase-lock/delay-lock circuitry. This choice of +15 dB for a nominal threshold is somewhat arbitrary, and could be pushed upward (or downward) by a few dB if we asked for an optimum balance between losses from synchronization errors and other factors. In Fig. 3, configurations of data rate times antenna size which are above or to the right of the indicated threshold lines will provide synchronization loop margin of greater than +15 dB for the indicated mechanization, and are thus acceptable to that mechanization. We recall that arraying requires alignment of both delay and phase, so that mechanisms which treat only delay (cases E or F) or only phase (cases A, B, D, or G) must be combined with a complementary mechanism to construct a complete system. The various mechanizations will be discussed in the following.

Case A

This is the carrier sync step of the standard DSN configuration. We use separate phase locked loops at each antenna of the array to perform carrier tracking and demodulation prior to combining at baseband. Assume that the spacecraft utilizes the approximate optimum modulation index as found by Greenhall (Ref. 1). Assume also that the topocentric doppler requires use of a 30 Hz tracking loop. For each data rate, the smallest acceptable single arrayable antenna is one which will just provide a +15 dB margin in the 30 Hz phase locked loop bandwidth.

Curve A also grants some insight into the carrier loop margin requirements: at data rates above about 3×10^3 bps, the whole array (single antenna) carrier margin is above the +15 dB margin level, while at lower data rates it is below, implying that the criterion of +15 dB loop margin for all synchronizing circuitry may be too stringent at low data rates, and too easy at high rates.

Case B

This is again the mechanism for carrier sync in the standard DSN arraying configuration, but in this case the spacecraft has

accommodated the fact of arraying by lowering its modulation index to provide a stronger carrier signal to be phase-locked at the individual antennas. The penalty, of course, is lowered data power. Suppose that the total array consists of several large antennas plus one small one. The smallest antenna's carrier margin will dominate the selection of mod index if we require that its margin be +15 dB or more. That small antenna would be *too small* to consider under the +15 dB carrier margin constraint if lowering the mod index to achieve that margin level also lowered the aggregate data power into the overall array by more than would be lost by simply abandoning the small antenna. This size limit to the smallest acceptable antenna is shown as Case B on Fig. 3.

Case C

Total spectrum combining. The combiner must track both differential carrier phase and envelope delay. Of these, the carrier phase is the more critical, and is assumed to be tracked in a bandwidth of 0.03 Hz. Use of this bandwidth for a 20 km baseline is approximately consistent with use of a 10 Hz bandwidth for the earth-rotation component of the full topocentric doppler. Both could potentially be reduced by a small amount. The envelope delay varies significantly more slowly, could be tracked with a loop of 0.003 Hz or smaller bandwidth, and thus would have a much higher loop SNR than would the phase tracking loop for the same conditions.

Assume that the array consists of M equal antennas. The input to the phase tracking loop is the time series of cross products of signal-plus-noise from each two of these M antennas. Let the data rate of the signal be B bits per second. With convolutional coding, like that of the Voyager, the channel rate is expanded to $2B$ symbols per second. If this symbol stream is modulated onto a subcarrier with N subcarrier cycles per symbol, and the resultant modulated onto a carrier, the spacecraft signal occupies a bandwidth of $2 \times (N + 1) \times 2B$ when only the first harmonics of the subcarrier are included, and several times that if higher harmonics are needed. For the purposes of estimating differential delay, we assert that the right signal bandwidth to use is one which just encloses the first harmonics. This filtering "gives away" about 1 dB of signal power in the subcarrier harmonics, plus another 1/2 dB in the upper-sideband data spectrum harmonics, while eliminating a much larger proportion of noise. The estimation process is a close analog of the Costas or squaring loop, and the performance is usually limited by noise times noise terms. Let the energy signal-to-noise ratios of the detected bit stream be R (actually 2.6 dB, or so) at threshold. Then the signal-to-noise ratio of the *bandpass* signal (first harmonic zone) from each of the M antennas is approximately:

$$\frac{R \times B}{M \times 2(N+1) \times 2B} \qquad \text{(less 1.5 dB)}$$

In forming differential phase/delay estimates, this bandpass signal is sampled, cross-products taken, and the resultant signal filtered to approximately a 0.03 Hz bandwidth. Assuming that the bandpass signal is dominated by noise, we find that the loop SNR of the phase tracking process is:

$$\text{Loop SNR} = \left(\frac{R}{M} \times \frac{B}{4(N+1)B}\right)^2 \frac{4(N+1) \times B}{0.03} \text{ (less 3 dB)}$$

$$= \frac{R^2}{M^2} \times \frac{B}{4(N+1) \times 0.03} \qquad \text{(less 3 dB)}$$

which we are insisting should be +15 dB. For $R = 1.8$ (2.6 dB), the phase tracking loop SNR is above +15 dB whenever

$$B \geqslant 2.2 \times (N+1) \times M^2$$

The case C line on Fig. 3 corresponds to $N = 10$ (subcarrier cycles per symbol), which is approximately the expected Voyager configuration at its Uranus or Neptune encounters.

Case D

Carrier combining. As in Case A, assume that the modulation index corresponds to that identified by Greenhall for the optimum for the aggregate array. As in Case C, the differential phase tracking is performed via filtered cross products of the signal as received at two antennas. Here, however, the predetection filter needs only to be wide enough to encompass the carrier, e.g., about 300 Hz. This 300 Hz bandwidth should be workable, but would probably require use of an ephemeris-tuned local oscillator. The tracking loop filter is again 0.03 Hz, providing a +40 dB increase in SNR between the raw cross products and the resultant phase-estimate. As we are demanding a +15 dB tracking loop SNR, we can accept a cross-product SNR of -25 dB, or a filtered signal SNR of -12.5 dB prior to the cross-product operation.

The available SNR for the 300 Hz prefilter for the whole array has been worked backwards assuming that the bit SNR is 2.6 dB, and using the approximate optimum mod index as shown in Table 1. The smallest acceptable single antenna size for carrier combining is that which will provide at least a -12.5 dB SNR in 300 Hz bandwidth for the individual antenna.

As an example, a data rate of 3×10^4 bps gives a carrier SNR in the 300 Hz predetection filter of +10.5 dB in the total array. An individual antenna with collecting aperture 23 dB below that of the total array would just provide a -12.5 dB SNR individually in that bandwidth, and thus be the smallest antenna acceptable. Other values are exhibited in the case D line of Fig. 3. We see that carrier combining can handle very

large arrays indeed (about 500 equal-sized antennas in the example case!).

Case E

Baseband combining. This is the second step of the current DSN arraying operations. This case is very similar to the total spectrum combining, Case C, but differs in two important considerations. First of all, the carrier phase information has been removed from the signal and does not need to be tracked – the remaining differential envelope delay is more slowly varying, and can be tracked in a narrower bandwidth, say 0.003 Hz. Secondly, the signal has been coherently folded about the carrier so that its bandwidth occupancy is lessened by a factor of two. Following the notation of Case C, the SNR of the pre-cross-product signal is

$$\frac{R}{M} \times \frac{B}{(N+1) \times 2B} \qquad \text{(less 1.5 dB)}$$

for which the loop SNR becomes

$$\frac{R^2}{M^2} \times \frac{B}{2\,(N+1) \times 0.003} \qquad \text{(less 3 dB)}$$

which should be +15 dB for satisfactory operation. The delay tracking loop is above +15 dB whenever

$$B \geqslant 0.11 \times (N+1) \times M^2$$

The case E line on Fig. 3 corresponds to $N = 10$.

The foregoing analysis assumes that the predetection filters indicated in Fig. 1 have been set to the near-ideal bandwidth for the combining process. If this filter bandwidth is made wider, or the filter eliminated entirely, the Case E line will move to the right in Fig. 3. For example, should the predetection filter be eliminated from the block diagram, and the baseband bandwidth left at approximately 3 MHz to cover the the first 9 harmonics of the Voyager subcarrier, the result would be to reduce the pre-cross-product SNR in the control path by a factor of about 8 (9 dB), and to move the Case E line on Fig. 3 *to the right* (larger antennas) by 9 dB.

Case F

Subcarrier demodulation and symbol detection before combining. The limiting conditions will be those imposed by the need for phase lock to the subcarrier. The loop bandwidth must accommodate both the topocentric doppler (scaled to subcarrier frequency) and modest oscillator variations, as well as limits to the human observers' patience in awaiting loop lock. If a 30 Hz loop can handle the X-band doppler, then a subcarrier loop of approximately 0.03 Hz should handle the doppler on the 360 kHz subcarrier of the Voyager. That value is used for both subcarrier tracking loop bandwidth and symbol synchronizer loop bandwidth in the following discussion.

A "squaring loop" provides the example configuration for a subcarrier demodulator. After coherent detection by the local carrier reference, the signal, consisting of data-modulated subcarrier plus noise, is bandpass-filtered about the fundamental of the subcarrier to remove excess noise and then squared to produce a pure unmodulated tone at twice the subcarrier frequency. This can be tracked to develop a coherent subcarrier reference which would be used to demodulate the data stream. The best bandpass filter will just enclose the fundamental zone of the data spectrum, thus giving away about 1 dB of signal power in the subcarrier harmonics, and another 1 dB in the harmonics of the data spectrum. Following the notation of Case C, the bandwidth occupancy of the signal is $4B$, and the SNR of the pre-squaring signal is

$$\frac{R}{M} \times \frac{B}{4B} \qquad \text{(less 2 dB)}$$

for which the SNR of a 0.03 Hz loop becomes

$$\frac{R^2}{M^2} \times \frac{B}{4 \times 0.03} \qquad \text{(less 4 dB)}$$

which should be above +15 dB for satisfactory operation. For $R = 2.6$ dB, the subcarrier tracking loop SNR is above +15 dB whenever

$$B \geqslant 2.9 \times M^2$$

This is shown as the Case F line in Fig. 3.

Case G

Data-aided carrier tracking at the individual antennas. In the best-of-all worlds, the detected data stream from the combined array is fed back to the individual antennas to allow the data modulation in the sidebands to be eliminated and the full power of the signal made available to be tracked as the carrier reference. There are no nonlinearities at low SNR to generate quadratic dependencies, so the loop margin in a 30 Hz loop becomes

$$\frac{R \times B}{M \times 30}$$

which, for $R = 2.6$ dB, is above +15 dB whenever

$$B \geqslant 500\,M$$

This is shown as the case G line on Fig. 3. This mechanism is more capable than tracking the residual carrier on individual antennas, but not as capable as carrier combining. There would also be some difficulty in initial acquisition of references needed to allow the decision feedback.

III. Two Examples

Two examples from potential arraying events should help to illustrate these results. The Voyager spacecraft at its Neptune encounter in 1989 could communicate at a data rate of 14.4 kbps if it is supported by an array consisting of all of Goldstone plus a large part of the VLA with masers for low-noise front-end amplifiers. The total array receiving aperture would be about five times that of the current DSN 64-meter antenna, while the 34-meter antenna at DSS 12 has a receiving aperture which is one-quarter that of the 64-meter: this antenna is thus 13 dB below the total array. The individual VLA antennas are expected to represent an aperture which is 18% of that of the 64-meter and are thus 14 dB below the size of the total array. Referring to Fig. 3, we could not reasonably expect to achieve 15 dB phase lock on the individual carrier signal at either of these antennas. We could, however, successfully involve either of them with total spectrum combining, or carrier combining coupled with baseband combining, and achieve this synchronization at the +15 dB loop SNR.

The second example is somewhat less encouraging. The Pioneer 10 and 11 are on their way out of the Solar System, and will be straining our ability to communicate with them by the end of this decade. At their extreme range, these spacecraft will be operating at S-band at 16 bps with an E_b/N_0 of 3 dB, and a fixed modulation index of 42°. These parameters represent a P_c/N_0 of 16 dB into the whole array, so that if we ask for a carrier margin on the whole signal (after arraying of the carrier) of +15 dB, we require the use of a 1 Hz phase-locked loop. At S-band, with no planetary bodies nearby to speed up the doppler, this narrow bandwidth could be workable. If we array carriers with a 300 Hz pre-detection filter and a differential loop bandwidth of 0.01 Hz, we get a +45 dB boost in SNR in going from the raw cross-products of the differential phase measurement to the filtered loop estimate of it. Asking for a +15 dB SNR in this differential phase tracking loop means that an SNR of -30 dB in the raw cross products is acceptable, and hence that an SNR of -15 dB in the 300 Hz pre-cross-product bandwidth is also acceptable at the individual antennas. At threshold conditions, the carrier SNR in 300 Hz for the whole array is -10 dB, which implies that the individual antennas should be no smaller than -5 dB, or one-third the total array. We could flex this limit moderately by narrowing the predetection bandwidth, or by accepting an arraying loop SNR of less than +15 dB, but even with these steps, it seems unlikely that an array like the VLA could be readily used to support the Pioneers at their extreme range.

IV. Discussion – Asymmetric Arrays

Although the preceding paragraphs are derived for a symmetric array where all antennas, including the reference element, are the same size, it is relatively routine to state similar results for asymmetric conditions: for Case A, B, and F, the antennas are treated individually, and the *smallest* one must be capable of supporting phase tracking at the specified SNR. For Cases C, D, and E, the phase and delay tracking are all performed using cross products of signals from separate antennas. If one antenna is larger (the reference antenna) but its bandpass is still noise dominated, then the SNR of the cross products is the product of the individual SNR's. For this case, we can interpret the behavior of the asymmetric array as if it were a symmetric array with the size of each apparent antenna being the square root of the product of the sizes of the actual reference and auxiliary antennas. In the limit, if we try to extract the ultimate performance from an array, we could attempt to use the aggregate array as the reference for each of the individual elements (Cases C, D, and E). Assuming that the acquisition problems can be solved, then the apparent performance boundaries for this case correspond merely to an asymmetric array with its reference antenna equal to the whole array.

Acquisition of the initial phase and delay estimates would not be easy, and would impose its own limit to operability of the array. One way to acquire the initial phase/delay estimates is to utilize one of the array elements as the reference antenna, achieving a +3 to +5 dB SNR, and then shift the reference to the aggregate signal to achieve low-loss signal combining and detection. This is a complexity which is probably best avoided, but would allow operating with smaller element antennas (or larger arrays). By accepting a phase acquisition SNR of only +5 dB, we can accept cross-product SNR's which are 10 dB less than those required for the Fig. 3 lines, and hence can accept single antenna SNR's or antenna sizes which are 5 dB smaller than those of Fig. 3.

V. Summary and Conclusions

There are limits to our ability to utilize ever larger arrays of modest-sized antennas to acquire data from distant spacecraft, but, as indicated on Fig. 3, they do not pose serious limits to data acquisition from Voyager or similar-capability spacecraft using the VLA. Trying to follow the Pioneers at a data rate of 16 bps out to the heliopause would be better done with an array of only a few larger antennas.

Acknowledgment

Many people have contributed over recent years to the ideas and material presented here. Notable among these are D. W. Brown, S. A. Butman, and W. J. Hurd of JPL, and P. Napier and A. R. Thompson of the VLA.

Reference

1. Greenhall, C. A., "A Compact Presentation of DSN Array Telemetry Performance," *TDA Progress Report 42-71*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 137-142, Nov. 15, 1982.

Table 1. Approximate optimum mod index for various data

Bit rate, bps	Mod index, deg	Carrier SNR/300 Hz, dB
1.2×10^5	80	+13.7
3×10^4	76	+10.9
10^4	71	+8.6
3×10^3	65	+6.0
10^3	59	+2.9
300	40	+0.0

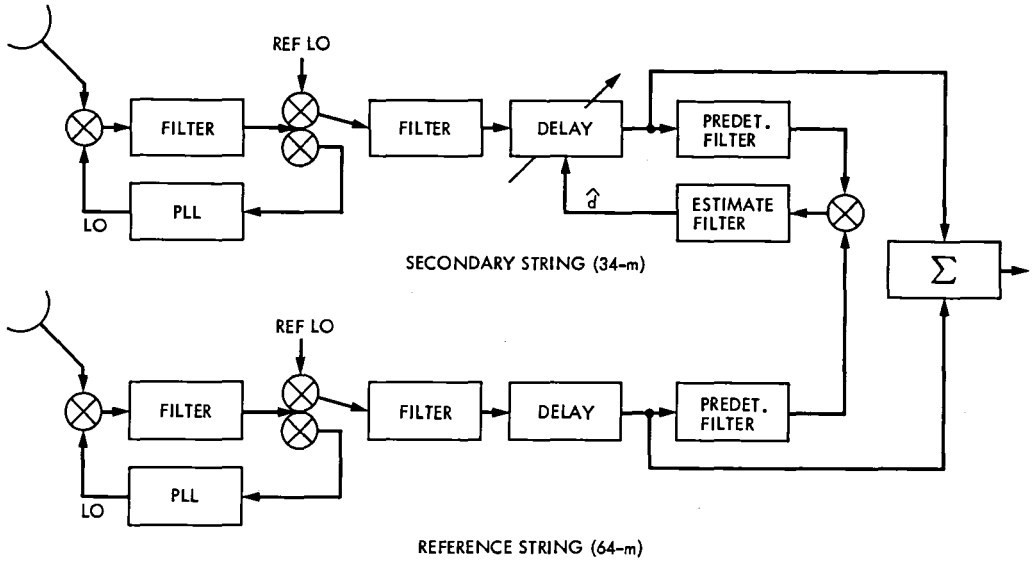


Fig. 1. DSN Combining process

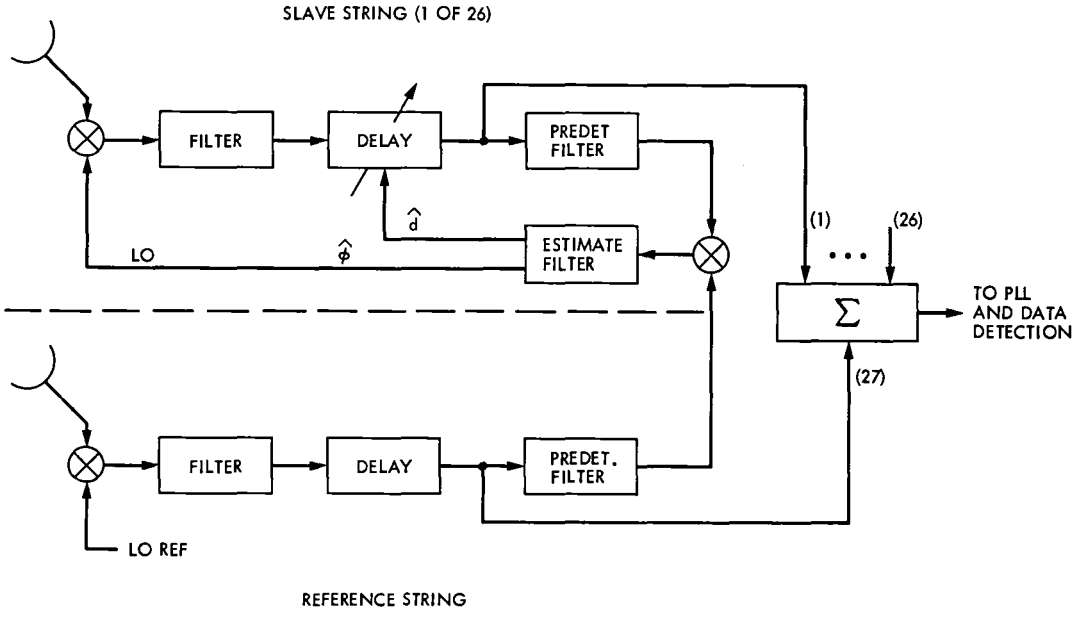


Fig. 2. VLA Combining process

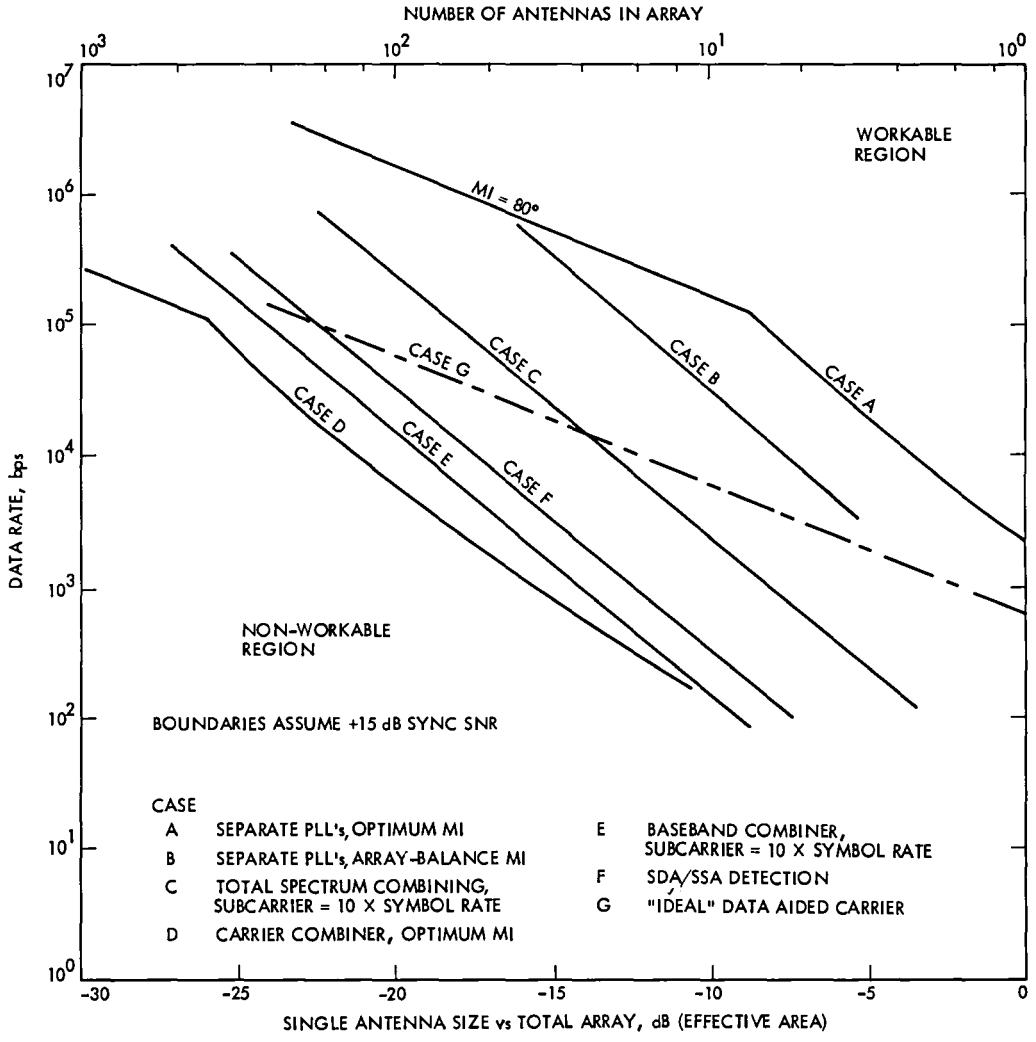


Fig. 3. Arraying limits

Developments Related to the Future Use of the 32-GHz Allocation for Deep Space Research

N. F. de Groot
TDA Technology Development

This paper was prepared for presentation at the third annual meeting of the Space Frequency Coordination Group (SFCG), held in Bangalore, India, at the invitation of the Indian Space Research Organization. SFCG is an informal organization consisting of the European Space Agency and agencies of 16 countries interested in space research. The purpose of the paper was to inform the SFCG of JPL/NASA activities that will lead to use of the 32-GHz allocations for deep space telecommunication. The paper discusses the advantage of 32-GHz links, the estimated link performance, and developments in the fields of antennas, low noise preamplifiers, radiometry, and propagation studies. A description of ways to demonstrate 32-GHz link capability concludes the discussion.

I. Introduction

In preparation for the 1979 WARC, the United States examined the need for new space research allocations. As a result of that examination it was proposed that a pair of bands near 32 GHz be allocated for space research, deep space only. At the WARC, 31.8-32.3 GHz was allocated for space-to-earth links, and 34.2-34.7 GHz for earth-to-space links.[1]

Operation at 32 GHz offers particular advantages, as compared to the lower frequency bands. To realize these advantages, there must be improvements in equipment and techniques. A substantial amount of analysis, design, and development is needed to achieve the improvements.

In this report we explain the advantages to be gained by operating at 32 GHz. Improvements needed to realize these advantages are identified. Several related study programs and hardware developments are then described.

II. Advantages of 32-GHz Links

There are two principal advantages offered by operation at 32 GHz as compared to lower frequencies: charged particles along the propagation path have less adverse effect, and link performance under clear weather conditions is generally higher.

For purposes of spacecraft navigation, and certain scientific measurements, it is important that the velocity of radio wave propagation be known very accurately. Interplanetary or atmospheric charged particles (electrons) alter the velocity to a degree proportional to the total number of particles along the

[1] These are footnote primary allocations that apply to Australia, Spain, and the United States.

path. To the extent that the number of particles is unknown or imperfectly modelled, the velocity will be correspondingly uncertain. The charged particle effect on propagation velocity is inversely proportional to the square of the frequency ratio. The velocity uncertainty at 32 GHz is thus 14 times smaller than that at 8.4 GHz, a clear advantage in favor of the higher frequency.

Another effect, particularly of dense plasmas (high concentrations of electrons), is the spectral broadening of signals passing through the plasma region. The broadening is noisy and limits the ability of an earth station to track weak radio signals that have passed through such a region. The sun is surrounded by an especially dense plasma. As the propagation path from a spacecraft in deep space comes closer and closer to the sun, the received signal on earth finally becomes too noisy for successful reception, and communication is interrupted. Operation at 32 GHz would allow reception of signals along paths passing closer to the sun, minimizing the time of interruption. This is important because the data obtained when a spacecraft enters and leaves occultation is of particular scientific value.

The second advantage to deep space links that is offered by 32 GHz is the potentially higher link performance, as compared to the 2- and 8-GHz bands currently used. The performance increase can be used to provide higher data rate or to allow communication using smaller and lighter antennas and transmitters.

III. Estimated Performance Advantage of 32 GHz

In free space, and with ideal fixed aperture antennas, link performance increases as the square of the frequency ratio. A 32-GHz link is thus 11.6 dB better than an 8.4-GHz link. Practical antennas, transmitters, receivers, and the propagation characteristics of the atmosphere reduce this potential advantage.

Table 1 presents the results of two analyses of the performance improvement that might actually be achieved.

The 1978 estimate is shown as historical background. It was made before the WARC, used data available at that time, and provided the justification for further development aimed at utilizing 32 GHz for deep space links. The 1982 estimate makes use of more recent weather data as described later in this paper.

Five factors were considered in each analysis: spacecraft efficiency, earth station efficiency, tropospheric loss, system noise temperature, and frequency. Spacecraft efficiency refers to the variation of e.i.r.p. caused by frequency-dependent changes in antenna gain and transmitter output power. Earth station efficiency refers to the performance of the 64-m antennas. Each factor was evaluated at 8.4 GHz and at 32 GHz. The ratio of the factors at the two frequencies gives the estimated relative gain or loss in performance that would be obtained by operation at 32 GHz.

For the analyses summarized in Table 1, spacecraft efficiency is assumed to be constant with frequency except for the traveling wave tube power amplifier. For deep space applications, the combined requirements of high efficiency and long life are particularly difficult goals to satisfy, and somewhat less efficiency is expected at 32 GHz.

The earth station efficiency is estimated on the basis of improvements to the existing 64-m antennas. Aperture phasing errors caused by the atmosphere are taken into account. For purposes of a conservative comparison, the ground antenna improvements assumed for the 32-GHz case are also applied to the 8.4-GHz case.

The third factor is the tropospheric loss that will not be exceeded at the Goldstone tracking station for 90% of the time. The fourth factor, system noise temperature, includes maser noise temperature and the estimated antenna temperature that will not be exceeded at Goldstone for 90% of the time. Cooled waveguide is assumed for the 32-GHz case.

The fifth factor is that of link improvement due to frequency increase under ideal conditions, as mentioned earlier.

IV. Comparison of the Two Analyses

The 1978 analysis in Table 1 shows a potential performance improvement between 5.8 and 1.4 dB. This unacceptably wide range of estimated performance improvement was principally caused by uncertainty in the tropospheric loss factor and its corresponding influence on system noise temperature.

To allow successful deep space missions using 32 GHz, it is necessary to substantially reduce the uncertainty in predicted performance. The alternative of providing excess link margin to cover the uncertainty is simply too expensive.

The 1982 estimate that is also shown in Table 1 includes more recent figures for tropospheric loss and system noise temperature. Not only is the uncertainty in estimated performance reduced from 4.4 to 1.2 dB, but the maximum potential benefit is higher than before.

V. Improvements Needed to Allow Deep Space Communications at 32 GHz

A number of improvements in earth station technique and equipment are necessary in order to realize the potential benefits of communicating at 32 GHz:

(1) Higher mechanical precision and pointing accuracy of earth station antennas.

(2) Wider bandwidth maser preamplifiers.

(3) More complete knowledge of 32-GHz attenuation and sky noise temperature at all three DSN (U.S. Deep Space Network) sites: Goldstone, Canberra, and Madrid.

Work on these is discussed in the following paragraphs.

A. Antenna Improvement

Earth stations of the U.S. Deep Space Network include 64-m reflector type antennas with Cassegrain feed. Table 2 shows their characteristics at 8.4 GHz. Designed 20 years ago, these antennas begin to lose efficiency at 10 GHz, and the gain limit is reached at 20 GHz. To achieve the estimated performance advantage of communication links at 32 GHz, it will be necessary to make a number of improvements to the antennas.

1. **Antenna precision.** The gain of a reflector antenna depends on its size and the degree to which the received wavefront can be collected and focussed without appreciable phase errors. There are several sources of these, and when the standard deviation of all the phase distortions approaches 1/12 wavelength, there is a 4-dB loss in gain. One-twelfth wavelength at 32 GHz is 0.78 mm. Since a 4-dB loss would appreciably reduce the advantage of 32-GHz links, it will be necessary to limit the phase distortions to the equivalent of a fraction of 1 mm.

Ideally, an antenna would have smooth and exact reflector and subreflector shapes, the correct alignment of these elements with respect to the feed, and would be so rigid that wind, gravity, and thermal effects could not disturb the total structure. In practice, this perfection does not exist, and the mechanical errors result in phase distortions that degrade the antenna performance.

Some phase distortions are caused by imperfections in the manufacture and installation of the mechanical parts of the antenna: the panels that make up the reflector, the subreflector, and the feed assembly. In addition, structural deflections caused by gravity and wind result in phase errors that are a function of elevation angle, pointing direction, and wind speed and direction.

To appreciate the difficulty of making an antenna that is suitable for 32 GHz, observe that the area of a 64-m reflector is approximately equal to one-half of a soccer field. Imagine keeping half a soccer field smooth and level to a tolerance of less than a millimeter!

2. **Antenna pointing.** In order to realize the gain offered by an excellent antenna, it is necessary to point it in the correct direction. The beamwidth of a given size of reflector antenna decreases with increasing frequency (up to the frequency where phase distortion limits the gain). The pointing at 32 GHz must be four times more accurate than the pointing at 8.4 GHz, because of the decreased bandwidth.

Antenna pointing error is the difference between the intended and actual direction of the electrical boresight. The pointing error is a composite of phasing errors and mechanical imperfection in positioning the antenna structure with respect to the desired direction. Referring again to the 8.4-GHz characteristics of the 64-m antenna (Table 2), notice that the beamwidth at 32 GHz will be 4 times smaller, or approximately 0.01 deg. The existing pointing accuracy of 0.02 deg must therefore be improved so that there will not be excessive loss during initial acquisition of spacecraft signals.

Improving the pointing accuracy of an existing large antenna is difficult and costly. One technique that circumvents this problem is to experimentally create a table of actual pointing direction as a function of mechanical position. To point in a particular direction, the antenna is then caused to take the corresponding position as read from the table.

Once a signal has been acquired, it is necessary to follow the movement of the spacecraft with respect to the earth station. Directional control by means of conical scan or monopulse techniques can result in electrical or mechanical beam steering that maximizes the received signal strength.

It is estimated that acquisition pointing accuracy of ±0.005 deg and active tracking accuracy of ±0.001 deg can be achieved with the 64-m antenna under calm or light wind (6-m/sec) conditions. An important reason for the possibility of such accuracy is that the 64-m antennas include a special optical alignment system. Called the master equatorial, the system closes azimuth and elevation servo drive loops in a way that ensures that the mechanical axis of the main reflector is driven to the desired direction. In this way, the imperfections in the drive and support mechanism do not introduce pointing errors. The master equatorial reference system is a key element in the ability to use the 64-m antennas at 32 GHz.

(Note: The foregoing discussion of antenna pointing is in the context of receiving signals from known spacecraft by a

process of acquisition and tracking. There also is a need to receive signals that are so weak as to be detectable only after minutes, or perhaps hours, of signal integration. In this case, active tracking is not possible. Echoes from planetary radar exploration, and signals from some distant natural sources, are examples. Pointing that is needed to satisfy these requirements is beyond the scope of this paper.)

3. **Performance of improved antennas.** Table 3 lists the factors discussed above along with estimates of the phase distortion and corresponding loss that could reasonably be achieved with improved antennas. Also shown is the loss associated with atmospheric turbulence that affects phase across the incoming wavefront.

The estimated performance depends upon a number of improvements to the existing 64-m antennas. Structural stiffening to reduce the deformation of the main reflector due to gravity and wind has been designed and installed on the Goldstone antenna. Improved panel manufacturing has been demonstrated. Based on work in the United Kingdom, improved methods for measurement of surface precision are being developed to allow more accurate panel setting. Motor-driven axial adjustment of subreflector position has been installed to allow focus correction for structural deflection as a function of elevation angle. Similarly, an adjustment to compensate for vertical beam shift is planned. A subreflector surface that could be slightly changed in shape is being considered as a further means of correcting for gravity-induced deformation of the main reflector surface.

Through a continual program of measurement, analysis, design, and implementation, the useful life and frequency range of the existing antennas at deep space earth stations can be extended, and the practical application of 32 GHz made possible.

4. **Clear aperture antenna development.** A more fundamental improvement in antenna performance is also under development. The typical reflector antenna with Cassegrain feed (Fig. 1) suffers from a loss of efficiency because the feed and the subreflector support structure block a part of the wavefront. Additionally, microwave thermal radiation from the surface of the earth is coupled into the antenna by reflection from these structures, increasing the system noise temperature. An offset feed and subreflector (Fig. 2) achieves a clear aperture.

The reflector and subreflector of a conventional Cassegrain antenna are circularly symmetric surfaces with approximately parabolic and hyperbolic curvature. An offset feed requires that the surfaces be modified accordingly in order to provide proper phasing. Modern analytic techniques allow the calculation of reflector and subreflector shapes that complement each other and result in unusually high antenna efficiency. A 1.5-m-diameter antenna has been constructed using these principles. Analysis and tests at 32 GHz show an aperture efficiency of 85%, believed by the designers to be a world record. The high efficiency is accompanied by sidelobe levels that are comparable to those of conventional antennas. By tapering the aperture illumination, unusually low sidelobes are possible by using the clear aperture principle.

B. Maser Improvements

1. **Receiver preamplifiers.** Maser preamplifiers have been used with multiple conversion superheterodyne receivers to provide system noise temperatures of 16 kelvins at 2.3 GHz and 24 K at 8.4 GHz. These downlink bands are relatively narrow, 10 and 50 MHz respectively, and are well served by 40 and 100 MHz, the bandwidths of masers operating at these frequencies. To take advantage of the higher link performance that is possible at 32 GHz, and to provide for the use of the 500-MHz allocation width, new masers are needed.

The gain-bandwidth product of a maser amplifier depends upon the nature of the maser material, the size and configuration for the maser structure, and the operating frequency. Operation of a given length of ruby maser material at 32 GHz would yield higher performance than the same length operating at 8 GHz. Compared to the bandwidths of existing 2- and 8-GHz amplifiers, a much wider bandwidth can be achieved in a 32-GHz amplifier of the same size.

2. **Maser operating cost and a new multiple frequency amplifier.** An important limitation to maser design is the need for cryogenic (very low temperature) cooling. The high gain and extremely low noise temperature of a maser amplifier are achieved by cooling it to a temperature (4.5 K) near absolute zero. The energy cost of providing this cooling for the 2- and 8-GHz maser, located at the DSN tracking sites, is substantial. Adding the 32-GHz capability would add more cost. Fortunately, the development of a wideband 32-GHz maser provides an opportunity to construct a new multiple-frequency low-noise amplifier that uses a single maser. This new amplifier is possible as a result of improvements in microwave solid-state upconverters.

An upconverter is a special kind of mixer. Utilizing a semiconductor diode, it transforms an incoming signal to a higher frequency by means of a local oscillator called a pump. The input and output frequencies differ by the pump frequency. A unique characteristic of the upconverter is the power gain associated with the frequency increase. Low noise operation comparable to a maser is possible with cryogenic cooling. The upconverter can also have a wide bandwidth.

Figure 3 shows a block diagram of a multifrequency low noise amplifier being developed for the Deep Space Network. A single 32-GHz maser is used for that band. 2.3- and 8.4-GHz upconverters provide for those bands by translating the input frequencies to 32 GHz for amplification by the maser. The bandwidth of the maser allows separate 2.3- and 8.4-GHz signals to be simultaneously amplified and then sent to different receivers for further amplification and detection.

It is expected that the multifrequency upconverter-maser amplifier will have a noise temperature of approximately 3 K and a bandwidth of 300-500 MHz.

C. 32-GHz Propagation Characteristics

Adverse weather in the form of clouds and rain can degrade or interrupt the reception of weak signals from deep space. Clouds and rain increase the attenuation and sky noise temperature, both of which reduce the signal-to-noise ratio. To successfully design and operate a telemetry link, it is necessary to know the effect of weather on link performance.

The 1978 estimate of performance, Table 1, included a consideration of weather effects at Goldstone, based on knowledge available at that time. Since then, a program of measurement and analysis has given a more complete understanding of atmospheric effects on signal-to-noise (SNR) that might be expected with 32-GHz operation at the Goldstone earth station. Future work at the Canberra and Madrid stations will complete the information that is needed for mission planning and operation.

1. Propagation measurements and modeling. To understand the relationship between weather and link performance, it is necessary to have a model of the statistical variation of attenuation and noise temperatures with time. It will then be possible to know the fraction of time that a particular SNR can be exceeded. Since attenuation and sky noise temperature are related by a simple expression, it is only necessary to measure sky noise temperature over time.

Although a model of 8.4-GHz noise temperature has existed for several years, it cannot be successfully extrapolated to 32 GHz. This is because the modelling errors tend to be multiplied by the square of the frequency ratio. For example, 1 K uncertainty at 8.4 GHz becomes a 14.2 K uncertainty at 32 GHz. Since this magnitude of uncertainty makes useful analysis and prediction impossible, it was necessary to make measurements directly at approximately 32 GHz, and this has been done.

Early in 1981, a water vapor radiometer at the Goldstone station measured the increase in sky noise temperature caused by the atmosphere. A simplified diagram of the instrument is given in Fig. 4. A square law detector produces a voltage proportional to the noise power received from the antenna or from hot and ambient temperature loads. These loads at known temperatures allow calibration of the radiometer and calculation of the sky noise temperature as seen by the antenna. Figure 5 is an example of the measurement made by a radiometer.

By correlating the radiometer data with other meteorological observations of temperature, pressure, humidity, rain rate, and cloud cover, it is possible to model the attenuation and noise temperature that may be expected over time. For example, at 32 GHz, the model indicates that the Goldstone sky noise temperature will be 29 K or less 90% of the time, at an elevation angle of 30 deg. Atmospheric attenuation is directly related to the sky noise temperature. Each 6.45 K of noise temperature indicates approximately 0.1 dB of attenuation.[2] The attenuation that corresponds to a 29 K sky noise temperature is therefore 0.45 dB.

It was mentioned earlier that 8.4-GHz models cannot be extrapolated to 32 GHz because of the manner in which errors are multiplied at the higher frequency. A 32-GHz model may, however, be applied to 8.4 GHz by calculation. Table 4 is an example of the steps involved. The computation uses the square of the frequency ratio with respect to the noise temperature contribution of water vapor, cloud, and rain. The contribution from the dry oxygen component of the atmosphere is drawn from other data.

Returning to Table 1, the 1982 estimate of performance advantage uses the new 32- and 8.4-GHz models of attenuation and noise temperature. The net advantage of operation at 32 GHz will equal or exceed a value of 5 to 6.2 dB at Goldstone for 90% of the time.

2. Prediction of link performance. We have seen how radiometer and other measurements are combined to permit a more complete understanding of the effects of weather on the performance of a 32-GHz link. By knowing the probability that a particular link degradation may take place, the spacecraft designers can plan how to meet mission requirements. Figure 6 shows an example of the predicted ratio of received signal power to noise spectral density as a function of time. Curves are given for each of the three earth station locations and for four different weather assumptions. (The curves in this example are for a particular 8.4-GHz telemetry link for the Galileo mission to Jupiter. The values shown may *not* be extrapolated to 32 GHz without serious error. A 32-GHz example of link prediction was not available for this paper.)

[2]This simple relationship is only valid for small (0 – 0.5 dB) values of attenuation.

Thus we see that radiometer measurements taken to improve the estimated telemetry performance advantage at 32 GHz may also be used as part of the data needed for more accurate navigation. Knowledge of atmospheric delay is also applicable to the 2- and 8-GHz bands used for current deep space missions.

VI. Working Toward a 32-GHz Capability

In the foregoing paragraphs we have discussed antenna improvements, maser development, and data collection for propagation models. These are just a beginning in the development of an operational capability at 32 GHz. The development will take several years and involve some important steps that will be described next.

A. Uplink Requirement for 32-GHz Downlinks

Coherent two-way communication is needed for the doppler tracking that is an essential part of spacecraft navigation. Two-way communication requires an uplink, a coherent frequency change in the spacecraft transponder, and a downlink. In principle, the current 2.1-GHz uplink transmitter could be used in conjunction with a 32-GHz downlink from a spacecraft. A 32/2.1 turnaround ratio would be required in the transponder. That ratio is too large and would result in excessive downlink phase jitter. Fortunately, a higher frequency uplink will be available in time for the 32-GHz requirement.

Uplink capability at 7.2 GHz is being developed for the Galileo mission to Jupiter. For a future spacecraft with a 32-GHz downlink, a 7.2-GHz uplink would allow an acceptable transponder turnaround ratio.

B. Capability Demonstration

In order to convince mission planners that 32-GHz downlinks are practical and offer advantages, it is necessary to demonstrate a reliable capability. In the absence of 32-GHz transmissions from spacecraft, another way must be found to show that the necessary receiving system development has been accomplished. There are three possibilities. First, a combination of an improved antenna and maser, plus the associated receiving equipment, could be operated over a protracted period without any received signal. Measurement of system noise temperature would allow demonstration of reliability and predicted performance. Correlation of the measured temperature with weather models would show the degree to which claimed performance advantage could be realized in practice.

Second, the receiving system capability could be demonstrated by listening to radio star emissions. Important data about antenna pointing would be gathered also.

Third, and most comprehensive, a 34-GHz radar experiment would provide valuable experience with both up and down links. The weak echoes from objects in space would test the receiving system. The radar experiment would also give experience with the equipment and techniques for needed high power transmissions in the 34-GHz deep space uplink band.

The eventual achievement of communication capability at 32 GHz is thus seen to be the result of laboratory analysis and development, field testing and experimentation, and careful planning to ensure that the necessary steps are taken according to a schedule that meets the future requirements of space exploration.

VII. Conclusion

In this paper we have seen that a 32-GHz telemetry link from a spacecraft in deep space could have a performance advantage of 5 to 6.3 dB, as compared to the 8.4-GHz band currently used. This advantage depends upon a number of improvements in equipment and technique. Some of these were discussed: antennas, low noise amplifiers, and propagation modelling. Ways of demonstrating readiness for 32-GHz operational links were also considered. In addition, we have seen how the study of needed 32-GHz improvements sometimes yields information that is also useful in the lower frequency bands.

The long and challenging process of developing the capability of using 32 GHz for deep space research has begun.

Table 1. Estimated 32-GHz performance advantage that will be equalled or exceeded for 90% of the time at Goldstone, as compared to 8.4 GHz

Parameter (See text)	1978 estimate			1982 estimate		
	Elevation angle, 30 deg			Elevation angle, 30 deg		
	Value at 8.4 GHz, X	Value at 32 GHz, K_A	K_A/X ratio, dB	Value at 8.4 GHz, X	Value at 32 GHz, K_A	K_A/X ratio, dB
Spacecraft efficiency	0.17	0.17 – 0.13	0.0 to –1.2	0.17	0.17 – 0.13	0.0 to –1.2
Ground efficiency	0.61	0.36	–2.3	0.61	0.36	–2.3
Tropospheric loss factor	0.97	0.85 – 0.69	–0.6 to –1.5	0.99	0.9	–0.4
System noise temperature, K	34 – 38	66 – 127	–2.9 to –5.2	24	45	–2.7
Frequency, GHz	8.45	32.0	+11.6	8.45	32.0	+1.6
Total, dB	–	–	+5.8 to +1.4	–	–	+6.2 to +5.0

Table 2. Characteristics of DSN 64-m antenna, 8.4 GHz

Gain	72.1 dBi
Beam width	0.036 deg (3 dB)
Polarization	RCP, LCP
Ellipticity	1.0 dB
Pointing accuracy	0.02 deg (master equatorial reference)
Pointing precision	0.002 deg
Efficiency	51%

Table 3. Estimated pathlength errors and associated losses for improved 64-m antenna

Error mechanism	Elevation angle: Zenith σ, mm	Zenith X-band loss, dB	Zenith K_a band loss, dB	30 deg σ, mm	30 deg X-band loss, dB	30 deg K_a band loss, dB	10 deg σ, mm	10 deg X-band loss, dB	10 deg K_a band loss, dB
Gravity (structure)	0.42	−0.10	−1.38	0.038	−0.00	−0.01	0.19	−0.02	−0.28
Wind, 32 km/hr (20 mph)	0.28	−0.04	−0.61	0.28	−0.04	−0.61	0.28	−0.04	−0.61
Subreflector manufacturing	0.25	−0.03	−0.49	0.25	−0.03	−0.49	0.25	−0.03	−0.49
Panel manufacturing	0.25	−0.03	−0.49	0.25	−0.03	−0.49	0.25	−0.03	−0.49
Panel setting	0.25	−0.03	−0.49	0.25	−0.03	−0.49	0.25	−0.03	−0.49
Tropospheric turbulence	0.12	−0.01	−0.11	0.18	−0.02	−0.25	0.30	−0.05	−0.70
Antenna pointing	0.15	−0.01	−0.17	0.15	−0.01	−0.17	0.15	−0.01	−0.17
Totals	0.69	−0.25	−3.74	0.57	−0.16	−2.51	0.64	−0.21	−3.23

Table 4. Example of 8.4-GHz sky noise temperature computation using data from 32-GHz model

	Parameter	Value K	Comment
(1)	Total 32-GHz tropospheric noise	29.0	From statistical model
(2)	Contribution from oxygen only	9.8	From other data for clear dry atmosphere
(3)	Contribution from water vapor, cloud, and rain	19.2	(1) − (2)
(4)	8.4-GHz contribution from water vapor, cloud, and rain	1.4	(3) ÷ $(32/8.4)^2$
(5)	8.4-GHz contribution from oxygen only	3.5	From other data for clear dry atmosphere
(6)	Total 8.4 GHz tropospheric noise	4.9	(4) + (5)

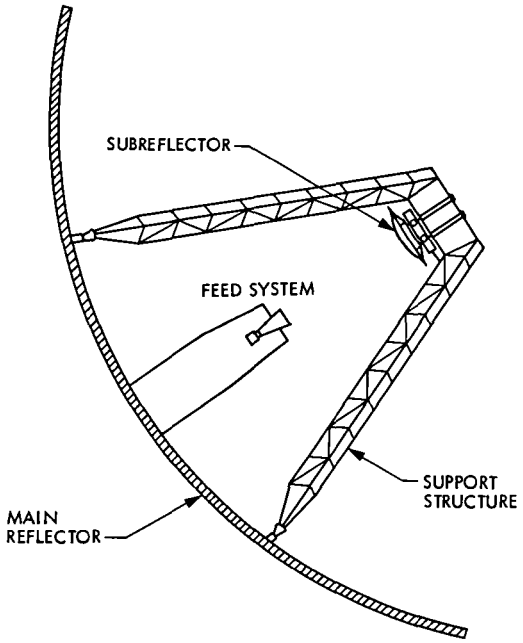


Fig. 1. Cassegrain fed dual reflector antenna

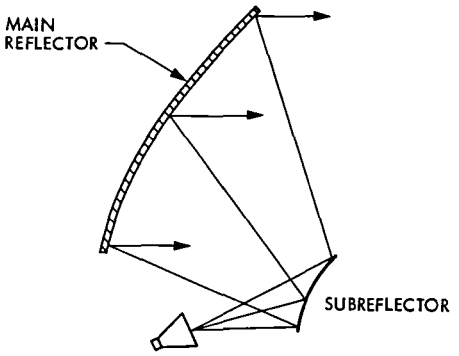


Fig. 2. Offset fed dual reflector antenna

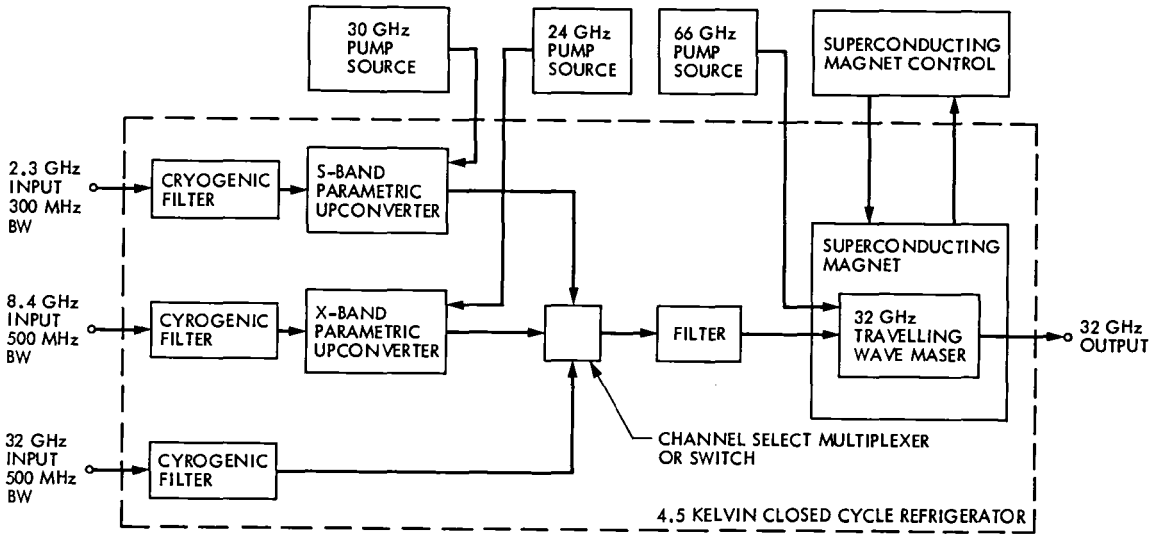


Fig. 3. Multifrequency low noise amplifier

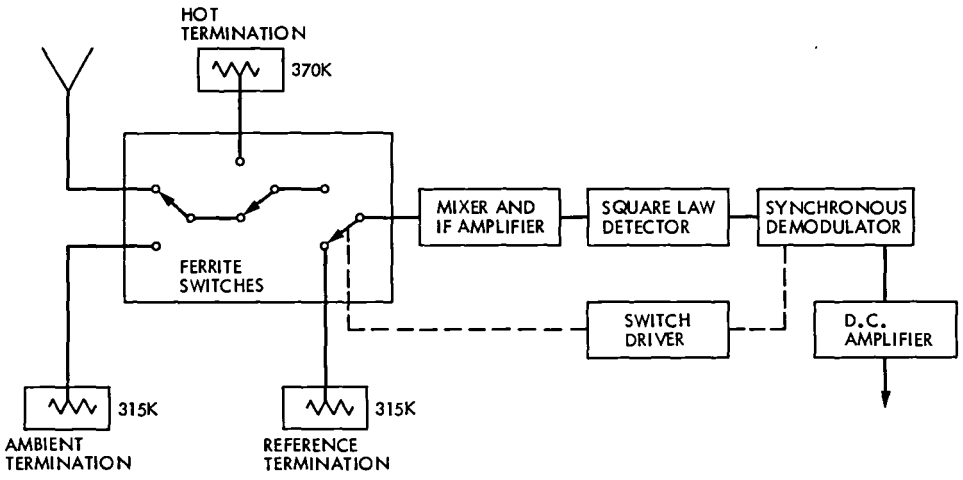


Fig. 4. Block diagram of radiometer

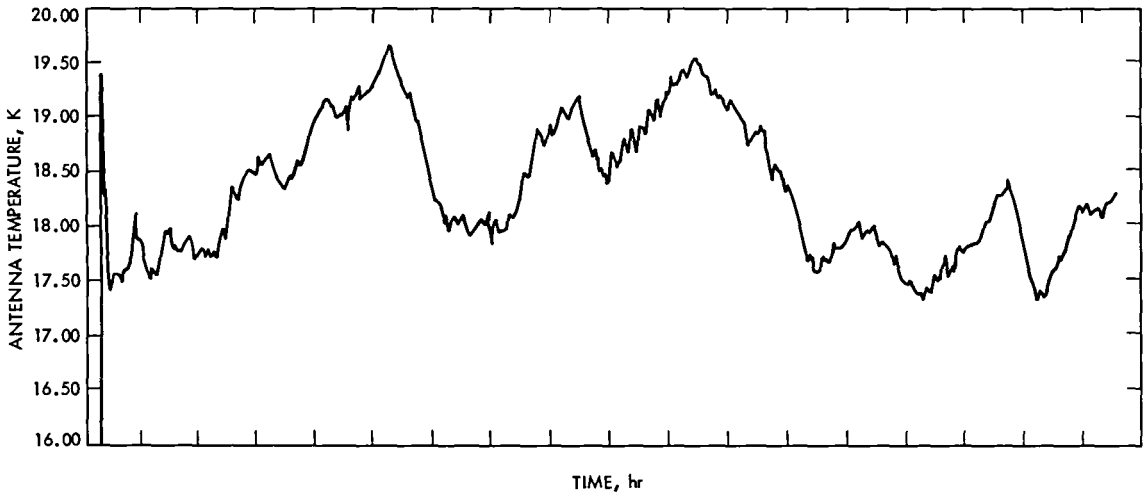
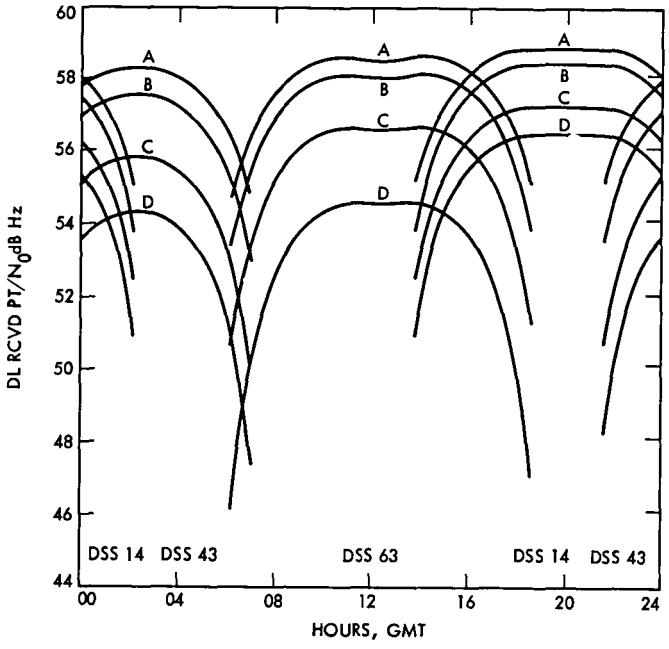


Fig. 5. Example of antenna temperature measurement made by radiometer at 31 GHz



A CLEAR DRY WEATHER, MEAN LINK

B STATISTICALLY DERIVED EXPECTED VALUES OF PERFORMANCE, COMBINED WEATHER AND LINK

C PERFORMANCE THAT WILL BE EXCEEDED 90% OF THE TIME, COMBINED WEATHER AND LINK

D EXPECTED VALUE OF LINK PERFORMANCE MINUS 2σ LINK TOLERANCE, MINUS 90th PERCENTILE WEATHER LOSS

PT TOTAL RECEIVED POWER, dBm

N_0 NOISE SPECTRAL DENSITY, dBm/Hz

DSS 14 GOLDSTONE, CALIFORNIA

DSS 63 CANBERRA, AUSTRALIA

DSS 43 MADRID, SPAIN

Fig. 6. Downlink PT/N_0 for Galileo mission vs time on 14 August 1991, 8.4 GHz

End of Document